Space and Time in Monoidal Categories

Pau Enrique Moliner

Doctor of Philosophy
Laboratory for Foundations of Computer Science
School of Informatics
University of Edinburgh
2020

Abstract

The use of categorical methods is becoming more prominent and successful in both physics and computer science. The basic idea is that objects of a category can represent systems, and morphisms can model the processes that transform those systems. We can see parts of computational protocols or physical processes as morphisms, which, when appropriately combined using tensor products and categorical composition, model the protocol or process as a whole. However, in doing so, some information about the protocols or processes is forgotten, namely in what location of spacetime did the events involved take place, and what was the causal structure among them. The goal of this thesis is to explore how these categorical models can be enhanced to include information on the spacetime location and causal structure of events.

First, we introduce the theory of subunits, which are subobjects of the monoidal unit for which a canonical isomorphism is invertible. They correspond to open subsets of a base topological space in categories such as those of sheaves or Hilbert modules, and under mild conditions they endow any monoidal category with a topological intuition. We introduce and study well-behaved notions of restriction, localisation, and support. Subunits in general form only a semilattice, but we develop universal constructions completing any monoidal category to one whose subunits universally form a lattice, preframe, or frame.

Afterwards, we introduce a number of constructions to explore how the theory of subunits can be used in practice. Inspired by logical clocks, we define a diagrammatic category where we can capture simple protocols and their causal structure. To progress towards more detailed spacetime and causal information, we define the category of protocols, which formalises the idea of letting a morphism from a category be supported in a different category. This allows us to have one category to model the systems and processes and another one to model spacetime. In particular, we can treat both toy models of spacetime and more realistic ones in the same mathematical footing. A notion of causal structure is defined for monoidal categories, and a generalisation of the usual causal analysis in physics for points to arbitrary regions is provided. We give examples of protocols seen as diagrams and as objects in the category of protocols, both with toy models of spacetime as well as with more realistic ones.

Acknowledgements

First and foremost, I would like to thank my principal supervisor, Chris Heunen. Since I was young, and even before being aware, I have been in a journey towards a greater conceptual understanding of science and mathematics. This brought me to study physics, mathematics, and some computer science and philosophy. Eventually, I found that a heavy focus on mathematics is good for clarity, but a strong intuitive influence from the other fields is very desirable. Fortunately, I discovered the field of categorical quantum mechanics, which in my opinion strikes a great balance between all these areas of knowledge, and in particular I am confident that Chris's approach resonates the most with the one I was looking for. So, I feel very lucky to have had the opportunity to work with Chris, and to have had him guiding my progress towards understanding and contributing to this interdisciplinary field. Furthermore, Chris is a role model for academic efficiency and achievement, which is really inspiring. To me, Chris has represented a lot more than can be hoped from a supervisor. I am very thankful for his constant help and support, in particular to find mental stability in the difficult moments, without which this thesis would not have been possible. I would also like to thank my secondary supervisor, Petros Wallden, for his interest in the project and for always being available for interesting discussions. Finally, my previous supervisors José Bordes, José Adolfo de Azcárraga, and Ingo Runkel have been key in guiding and helping my growth, and I am therefore indebted to them.

I thank Sean Tull, Pablo Andrés Martínez and Rui Soares Barbosa for the interesting discussions, as well as those I met in my research visit to Birmingham in February 2019: Steve Vickers, Jamie Vicary, Noam Zeilberger and Nicolas Blanco. They have directly contributed to my understanding of some aspect of this project, and to my motivation to try different things out and to learn more.

There is an increasing number of people who understand the potential of categorical semantics and want to contribute to making it central to our understanding of the world. This is demonstrated by the growing number of publications, conferences and journals on the connection between category theory and science, see for instance [17], [82], [29], [43] and [21]. I feel lucky to have been involved in research during the early growth of this community, in which so many points of view and areas of knowledge are beautifully combined. I hope that this community will benefit from the ideas in this thesis.

Another group I am indebted to is the Basic Research Community for Physics (BRCP), of which I became a member in 2016, since its passion to provide more clarity in science and its foundations has been truly inspiring. In particular, the BRCP-organised workshop Rethinking Foundations of Physics 2017 was really eye-opening as to the potential for interdisciplinary collaboration, and it was overall the best workshop I have ever attended. I thank fellow members Robin, Johannes, Federico, Joanna, Pierre and Sina, among others, for interesting discussions.

My parents and my sister have always supported me, which has allowed me to freely engage in creative and intellectual pursuits (a luxury that most of our ancestors never had). They always believed in me, and made me think that I am capable of achieving any goal I set for myself, provided I am perseverant. It is thanks to them that I developed my passion for education, which I see as the main key to the progress of both individuals and societies. This project, or any of the major achievements in my life, would have not been possible without my family's support, and for that I will always be grateful.

A special mention goes to the good friends I made during my childhood and which I still preserve today: Fran, José Ramón, José Francisco, Juamba, Alba and Rober. They have all helped me stay close to my Valencian roots, and have kept the visits to Spain exciting thanks in particular to their brilliant sense of humour, which is invaluable to me. I am thankful that they have stayed in touch despite the distance (in space and time). From my times at various universities, I would like to mention Antonio, Oscar, Héctor, Álvaro, Wadim, Lóránt, Lukas and Áron, to which I thank for sharing their passion for science and maths, as well as many memorable moments. My Erasmus visit to the University of Leeds was crucial in my academic development, and I would like to thank Alberto and Irene, among the many other friends I made there, not only for making the year really enjoyable, but also for helping me make it a great academic success. Finally, I thank Kiran for helping me feel at home in UK and for the direct support during the thesis writing.

Carlos has been a major influence since we met when starting our undergraduate studies at the University of Valencia. I am indebted to our many conversations, collaborations and mutual challenges, since they have shaped my way of thinking into a more mature and aware one, which in turn has been crucial in making me a more independent thinker. I thank him for sharing a pursuit of beauty through mathematics, music and strategic games, as well as for putting his energy in the LEFT (currently SEMF) society we co-founded, and I wish for us to share many more projects in the future.

Above all, I want to thank Chahat. She has allowed me to discover a a great deal about what I really want in life, and that it is possible to balance being hard-working

while remembering that life goes well beyond career. Since we met, Chahat has been an inspiration on how to keep achieving and improving oneself while being kind and caring for all other beings. Her constant support and patience have been crucial to the completion of this thesis. My current identity and world view, as well as my improved capacity to stay calm and enjoy life despite the adversities, cannot be understood without her influence over the last few years. For this, and for many other reasons, I am forever grateful to her.

Lay Summary

Mathematics is often referred to as the language of the universe. What does this really mean? When we write, we encode concepts as words, which can be decoded back into concepts with a knowledge of vocabulary and grammar. Similarly, the musical language allows us to translate between emotions and sounds. Following this analogy, mathematics is a language in which we can make our intuitions about the world precise. This allows us to not only communicate about them, but also to make theories and predictions that would otherwise not have been possible, and this is crucial for both science and technology.

As a language, mathematics is in constant expansion and evolution. For instance, the mathematics that Newton used are very different than those Einstein had access to. Roughly, this means that more ‘words’ and ‘grammar’ are added to the language, thus allowing for more expressive power and often for more abstraction and clarity. More precisely, mathematics expands and evolves when we formalise new intuitions as definitions, and when we formally study their properties and how they relate to one another.

This is an interesting time in science, since fully understanding quantum physics and its potential application for quantum computers is a big unresolved problem. For instance, it is not clear how to achieve an integrated understanding of gravity and of quantum physics. The recent history of the field of algebra gives us an important example of expansion of evolution of mathematics towards more conceptual clarity, and our plan is to take advantage of this for understanding physics and its applications in computing. In particular, models for physical and computational processes based on so-called category theory have recently emerged, which, thanks to their abstraction, can focus on the key ideas while forgetting about the irrelevant details. As a result, categorical models can provide a more natural way to talk and reason about these processes. However, at the moment these models are limited to describing what happens in the processes, and they forget where and when. Our main goal is to expand this branch of mathematics so that we are able to also talk about the locations of events.

Because the speed of light is the maximum at which any information can travel, it may be impossible for some events to be causally related to some others. As a secondary goal for this thesis, we aim to explore the capabilities of categorical models to include information about causal relationships.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

The author's contributions to [26] and [27] include the notions and results on firmness, basic results and examples of subunits, restriction, localisation, graded monads, and the completions, especially the Day convolution aspects.

(*Pau Enrique Moliner*)

Table of Contents

Chapter 1

Introduction

1.1 Motivation and goals

Categorical methods

When trying to make sense of the world around us, abstraction is the main tool at our disposal. It allows us to identify which concepts should be considered as fundamental, and to place them at the core of a more elegant, general and far-reaching formalism. In fact, most changes of paradigm towards a better understanding of nature have, throughout history, been closely related to an increase in the level of abstraction in our way of thinking. Despite this, changes in point of view often find resistance from the scientific community. An important example is group theory: mathematicians considered it to be too abstract and without application, and physicists were reluctant to accept arguments based on this formalism [4]; however, now it is central to modern physics among many other applications. I believe that category theory can become another instance of a formalism which, although once considered to be too abstract, becomes central to our understanding of the world.

One of the main virtues of category theory is that concepts, constructions or arguments that are intuitively analogous in essence can often be formalised as separate instances of a more general definition, and such concepts can be understood better all at once by studying the general definition. In this way, entire fields can draw from each other, and a better understanding of the connections between them emerges. This is similar to what happened in the transition from classical algebra, which was focused on studying polynomial equations, to modern algebra. However, category theory goes beyond this in that it can be regarded as a foundation for mathematics, alternative to the traditional set-theoretic one (see [64] for a mathematical presentation, and [59] for a more philosophical one). Taking categories as the foundation makes it more natural to think in terms of universal properties, which not only define an object but also capture one of its essential properties. By contrast, set theoretic foundations force concepts to be encoded as sets, regardless of whether this is natural or essential to the concept (for example, each number is formally a set). Although thinking in categorical terms may not be the most practical in every situation, it is at least helpful to understand the bigger picture and often to provide more conceptual clarity.

Categorical methods have been widely recognised as crucial in theoretical computer science since the eighties (see [75] for an overview, and references therein such as [34]). In particular, they give a natural way to reason about compositionality, abstraction and representation-independence, so they turn out to be an excellent setting for denotational semantics. Category theory is strongly connected to type theory [46], which can in turn be seen as a foundation for functional programming [85]. In relation to this connection, it is worth mentioning the programme of univalent foundations of mathematics [87], which incorporates intuitions from computer science and can result in more practical foundations, in the sense that concepts or statements that are used in practice, such as treating isomorphic objects as equivalent, become more natural and formalised.

In the case of physics, the main example of categorical approach is Categorical Quantum Mechanics (CQM), which abstracts away from Hilbert spaces to monoidal categories with various operationally motivated properties [21, 43]. The result is an abstract theory of systems (objects) and processes (morphisms), of which quantum physics can be seen as a particular case. Composition of morphisms is interpreted as performing the processes in sequence, while the tensor product is thought of as parallel execution. In some sense, we can see the goal of this approach as seeing how much can be understood by essentially only assuming that systems and processes exist which can be composed.

One way to improve our understanding of the world around us is to move away from concepts or assumptions that are not compatible with our best theories or experiments, such as the thinking of heat as a substance, or the existence of the Ether. But it seems like, in some sense, we cannot get rid of the concept of composition. As argued in [69], because our brains have a limited memory, the only way in which we can make sense of complex things is by combining together our understanding of the simpler parts involved. For example, as a collective we only accept arguments and proofs as valid if they are constructed by composing simpler ones, even if individuals can experience thought or intuition differently. Another prime example is natural language, in which the meaning of complex sentences are a result of composing the meanings of their parts. Therefore, it is natural to take composition as a fundamental concept, and the best way to do this is by using category theory; for, as expressed in [69], the essence of categories is composition, and the essence of composition yields the definition of a category. As an interesting consequence of this choice, the mathematical formalism of CQM is also successful for natural language modelling and processing [53, 74].

In principle, there are many ways in which we could address our goals, which we introduce below, with concrete constructions. However, we are motivated to take an approach based on categorical thinking. Therefore, our primary focus will be to understand how to formalise, within the context of monoidal categories, the concepts that would be common to all of these approaches. We will introduce some of these more concrete constructions afterwards, with the advantage that they can now formally be seen as just different ways to feature the same general definitions.

Spacetime

The programme of applying category theory to physics started in 2004, when Abramsky and Coecke [2] realised that quantum theory does not depend on the details of Hilbert spaces or even linear algebra to work, but instead makes sense in more general categories. Despite its youth, it has led to straightforward derivations of many standard results (see for instance [43], [21], and references therein). However, it is lacking a natural way to reason about regions of space and time, and about where and when events may take place. This is often not a problem; on the contrary, if we only are concerned with how the information evolves and gets transformed throughout the computation, the location of events becomes an irrelevant detail and it is best to abstract away from it. However, in some instances it may be very useful, or maybe even necessary, to keep track of the locations of events. The main goal of this thesis is to give theory and applications for categorically reasoning about spacetime locations.

We introduce the theory of subunits, which represent a notion of space inherent in monoidal categories. This is because in our leading examples, such as categories of sheaves and of Hilbert modules, the subunits correspond to the opens of the base topological space, and because in general subunits form a semilattice. While developing the theory, we will primarily think of subunits as open subsets, which means we may not have clear access to points. Thus, in some sense the theory of subunits resembles a categorical version of point-free topology [90], in which mentioning points is avoided. Subunits will relate to morphisms via a notion of support, which will allow us to talk about where can processes happen in space and time. In general, these concepts and constructions work abstractly beyond the intuitions that originally guided them, but in Chapter 3 we show how they can be used to serve those intuitions.

Causality

The speed at which information can travel through spacetime is limited to the speed of light, which means it may be impossible for some events to causally influence some others. The *future (light) cone* or *causal future* of an event p refers to the set of points that can be influenced by it, and can also be understood as the collection of trajectories that move into the future of p with a speed never greater than the speed of light. Similarly, the *past (light) cone* or the *causal past* of an event p is the region of spacetime that contains all the points that could in principle have influenced it.

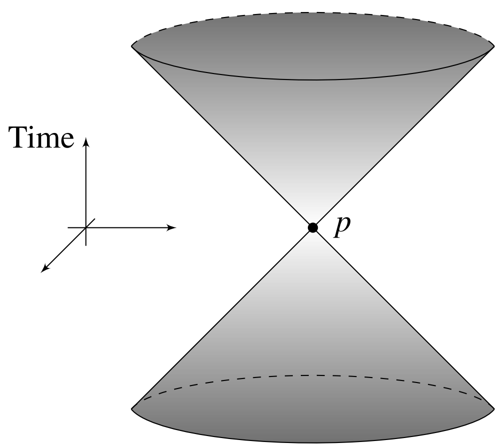


Figure 1.1: Light cone

In physics, spacetime is modelled as a Lorentzian manifold X with a time orientation [72], which we will call a *spacetime* manifold. In this setting, we can treat the causal relations formally: for points $s, t \in X$, write $s \prec t$ when there is a future-directed non-spacelike curve from s to t; in other words, when s can causally influence t. Then, we can define the future and past cones respectively as follows.

$$J^+(t) = \{s \in X \mid t \prec s\}$$
$$J^-(t) = \{s \in X \mid s \prec t\}$$

Interestingly, this causal relation allows to recover the manifold structure, under the mild physical condition of global hyperbolicity[1], by using domain-theoretic techniques [67]. For mathematical convenience, we can define a causal partial order $\preceq$ via $s \preceq t \iff (s = t \vee s \prec t)$.

There is some philosophical debate as to whether or not causality should be considered to be a fundamental concept in science, as some philosophers such as Russell argued that it should not [80]. One of the main reasons for the skepticism around the idea of causality has been that it is a concept too vague and woolly to be appropriate for physics or science. Thus, the skeptics argue, as science improves it should move away from this concept. However, there has recently been a great improvement on our collective understanding of the concept, for instance thanks to the work of Judea Pearl [25], and causality has become a precise mathematical theory with many significant applications [65, 30]. In particular, causal reasoning is a key tool for analysis and design in the study of distributed algorithms [60, Chapter 3]. I believe that causal reasoning is, at the very least, useful in science, and that it will continue to exist both in philosophy and in science (as it has done despite the Russelian pleads to the contrary). Therefore, I think it makes sense to work on and improve our understanding of causality, rather than trying to get rid of it.

In physics, events are assumed to happen in a single point of spacetime (in particular, instantaneously), and therefore causal structure is studied in terms of points. This thesis provides a way to lift this analysis from points to arbitrary regions, and gives a way to talk about causality in the context of the theory of subunits. The idea is that we can define future and past cones of regions as the unions of the corresponding cones of their points. Unlike in the case of points, causal structures for regions are separate for future and past, and in general cannot be induced from one another. Thus, when lifting the causal relations to regions, we get a pair of preorders, one for the future and one for the past, instead of a single partial order.

[1]According to Penrose [91], these are the physically reasonable spacetimes.

1.2 Structure of the thesis

The thesis is structured in two main parts. First, Chapter 2 introduces the theory of subunits. This contains most of the technical developments, but emphasis has been kept to clarify the intuitions that guided our definitions as well as the interpretations of the main results. Chapter 3, despite also containing abstract definitions and constructions, is primarily focused on exemplifying how the theory of subunits can be used towards accomplishing our goals as set above.

Chapter 2 is based on the contents of [26] and [27], whereas Chapter 3 (except Definition 3.22 and Proposition 3.23, which belong to [26], and Section 3.1) consists of new unpublished content that I developed on my own.

Thoughts on future developments have been split in two parts. At the end of the introduction, more speculative further directions are discussed, since they do not require technical understanding of the contents of the thesis and can serve as motivation. Chapter 4 focuses on those further developments that are more of a continuation of the thesis; in other words, the questions I would like to look at if time allowed.

Theory

More specifically, Section 2.1 defines subunits as subobjects of the tensor unit with an idempotence property, and shows that, with the small caveat of firmness, they always form a semilattice. Next, Section 2.2 characterises the subunits for different examples. In particular, subunits of the category $\mathbf{Hilb}_{C_0(X)}$ of Hilbert modules over a base topological space X correspond exactly to the open subsets of X. This gives us the interpretation of subunits as regions of space and time, and it can be seen as a generalisation to the non-cartesian case of the well-known fact that subobjects of the tensor unit in the category of sheaves $\mathbf{Sh}$(X) correspond to the open subsets of X.

The notion of restriction to a subunit is introduced in Section 2.3, as the result of lifting the corresponding topological intuition as well as the algebraic notion of restriction of scalars. For every subunit, the category can be restricted functorially to a coreflective monoidal subcategory. The family of all restriction functors forms a graded monad. Finally, and corresponding to these two results, alternative characterisations of subunits are given: one in terms of monocoreflective tensor ideals, and one in terms of idempotent comonads.

Section 2.4 shows that restriction is an example of algebraic localisation, and that localising all the subunits at once gives a universal way to make a category simple; that is, to obtain a new version of the category with trivial subunits.

Next, in Section 2.5 we interpret a morphism restricting to a subunit as meaning that the process it represents has support in the region the subunit models. The map that takes a morphism and returns the set of subunits in which it is supported is formalised as a functor in the definition of support datum, and we show that any firm category has a canonical support datum that can be characterised via universal property.

In the rest of the chapter, we work towards answering the following questions. When do the subunits have further structure than being a semilattice, such as being a lattice, preframe, or a frame? Can we find a construction that, given a category, completes its subunits to one of these further structures?

We will call *locale-based* those categories whose subunits form a frame, and which satisfy a technical condition of cooperation between subunits and the other morphisms that makes completing to a locale-based category always possible. Locale-based categories are defined and characterised in Sections 2.6 and 2.7, and we use this to show that our main examples are locale-based. Analogous to the requirement of being locale-based, we have slightly different (weaker) technical requirements when we only seek completion to a lattice or a preframe.

For each of the structures we want to complete the subunits to, we identify a relevant subcategory of presheaves under Day convolution (introduced in Appendix A). When completing to a locale-based category, we call the relevant presheaves *broad*. Sections 2.8 and 2.9 show that Yoneda embedding into the relevant subcategory of presheaves gives a universal way to complete; but this is not a sheafification for any Grothendieck topology.

Protocols

The idea behind this chapter is to exemplify how to revisit the categorical modelling of (distributed) protocols in order to add more information on their causal structure and on the spacetime location of the events involved. Examples will be presented roughly ordered by the increasing degree of complexity of the information they add to the model.

First, since many ideas in this chapter are inspired by logical clocks, these are briefly introduced in Section 3.1. This leads us to the category of diagrams over a given category $\mathbf{C}$, as defined in Section 3.2, which is a first attempt at adapting the idea of logical clocks to a categorical setting. While this succeeds in capturing the causal structure of protocols, it is not clear how to include more detailed information on the location of events, and it is also not clear how to relate it to the theory of subunits.

To address both of these issues, the category of protocols is introduced in Section 3.3. In it, protocols are seen as preorders of *events*, each event containing information on what happens and where in spacetime does it happen, while the preorder informs about causal structure. In practice, it may be very hard to have, in the same category, interesting morphisms (to model the possible processes) and interesting subunits (to model spacetime), so the category of protocols allows for different categories to each hold the relevant data. Its monoidal structure is defined so that it has meaningful subunits and interesting notions of restriction and support. Given a protocol, we can forget what happens on the events to obtain its *network*: the collection of all the locations involved in it. This can be seen as a forgetful functor, adjoint to the free functor that assigns the trivial process to every location of a given network.

Although spacetime is usually represented as a manifold in physics, we can have versions of it that are less 'realistic' and more like a toy model. Since we can now have a category dedicated to just having its subunits represent spacetime, we can for instance take any notion of spacetime that forms a semilattice and see it as a category. This is done in Section 3.4 for both a toy model based on logical clocks, as well as for a realistic model based on manifolds.

Next, Section 3.5 shows that the category of diagrams is equivalent in expressive power to the category of protocols when we limit ourselves to toy models of spacetime. In order to illustrate concretely some ways to use the category of protocols with toy spacetimes, we present a musical example and finally we show how to view quantum circuits as objects.

A discussion of causality in the context of subunits is given in Section 3.6. This is an interlude in the sense that the definitions and results introduced in it make sense in the general theory of subunits, regardless of what particular construction we may be using for applications; but we introduce them now since we do not use them before in the thesis. The operations of taking a region and returning its causal future or past are formalised as closure operators, and these are shown to be equivalent to preorders satisfying some intuitive axioms. This covers the transition from points to regions in causal analysis.

Finally, the chapter ends with examples of protocols with more realistic spacetimes, in Section 3.7. We build up, via a discussion on teleportation, to talking about the task of summoning information in spacetime. This demonstrates that we can indeed use our framework to categorically model protocols for which locations are essential. Viewing summoning as an object in the category of protocols, we formulate an argument that allows us to reduce the proof of the no-summoning theorem, which characterises the configurations in which summoning quantum information in spacetime is possible, to the categorical version of no-cloning.

1.3 Related work

Here we present a brief review of the literature that is most relevant to our project, with the goal of justifying why we take our own novel approach towards the aims described above rather than joining an existing effort.

Regarding the project of endowing CQM with the potential to model the spacetime location of events, we are aware of only one piece of related work, namely the paper by Blute and Comeau on Von Neumann categories [11]. They propose that an appropriate modification of algebraic quantum field theory would allow for such spacetime modelling, and with this motivation they reach the definition of a premonoidal quantum field theory as the basic setting with which to extend CQM. Using the Grothendieck construction to encode quantum protocols, they are able to model the quantum teleportation protocol analogously to the conventional CQM way in [2], but also accounting for the spacetime location of events. The main advantages of our approach with respect to this one is that we do not need tools outside of monoidal categories, and we require far less assumptions and elaborate constructions. Another difference is that they suggest to capture causality as a property via the bifunctoriality equation, which does not result in a mathematical structure for causality that one could work with and investigate further.

Categorical methods have been more successful at capturing causality than at capturing locations of spacetime. This is reasonable, since more information and detail need to be taken into account for the latter, even if the goal is to only have a toy model of spacetime.

The earliest work we consider for categorical treatments of causality is that by Coecke and Lal in [61, 22]. In it, causality is encoded as the possibility of information flow. From physical considerations, they derive the structural constraints that a symmetric monoidal category should satisfy in order to only represent processes that respect causal structure. This leads to the technical definition of causal category. Since this approach is based on restricting the categories typically used in CQM, causal categories are incompatible with several structures that are important in CQM, such as compact structure, the dagger functor, and monoidal products of the form $A \otimes A$. These can only be represented partially and indirectly in causal categories. As a result, it is not possible to find natural examples of causal categories, and the way to construct examples is quite intricate.

The main framework for causality modelling in CQM in the present can be found in [21, 56, 20, 57], and it is based in the principle of terminality. A symmetric monoidal category (or process theory) is called terminal if discarding the output of any of its morphisms is equivalent to simply discarding its input. This is argued to be necessary for science, because it means that, when doing an experiment, we can ignore processes whose outputs never reach the experiment before it ends. A process theory is said to respect causal structure if the output of a process only depends on what happened in the causal past of that process. As shown in [56], a process theory respects causal structure if and only if it is terminal. Terminality yields several other interesting results for general process theories (see Chapter 6 of [21]): relativistic covariance, unitarity for reversible pure processes and Stinespring dilation for general processes. Interestingly, this framework can be generalised in order to model processes which exhibit indefinite causal structure, as shown in [57].

Although a number of interesting results may be found arising from the principle of terminality, we think that a more structural and constructive approach, like the one we propose, is better suited to provide new directions of research and applications. In particular, our framework is more expressive, in the sense that we can for example refer to causal cones and their intersections. Finally, the existing frameworks for causality in CQM do not address the modelling of the spacetime locations of events.

Another interesting element of our treatment of causality is that, as explained in the previous section, we look at results for regions rather than just points. We think this is new, both in the categorical and in the broader mathematical treatment of causality.

This thesis is inspired by the idea of considering events as a primary concept, which can be found in both logical clocks (as introduced in Section 3.1) and in the theory of event structures [93, 92]. Event structures are a model of computational processes in which a set of events is considered together with relations that inform about how they may causally depend on each other, and how they may be incompatible or at conflict with each other. As explained by Winskel [92], and as is common in physics, what to call an event depends on the wanted degree of abstraction and detail in the model. The definition of the category of protocols in Section 3.3 is inspired by the formalism of event structures in that it consists of a set of events together with a relation $\sqsubseteq$ which has to do with its causal structure. Our ideas in Sections 3.4 and 3.5.1, where labels for location and time are formalised as subunits with which we can timestamp morphisms, are the result of adapting the idea of logical clocks to the theory of subunits.

Although there are categorical constructions in the theory of event structures (see for instance [93, Chapter 2]), our approach is different in that our central concept of subunits is already underlying in monoidal categories. It seems that event structures capture more naturally the incompatibility or conflict between events, although this could perhaps be added to our approach. In general, due to lack of time we were unable to formalise and explore further how our formalism may relate to the theory of event structures, so this has been left for future work (see Chapter 4).

1.4 Future directions

Relativity

When discussing protocols which are distributed among a number of parties, we will assume for simplicity that there is a priviledged frame of reference, so that coordinates have a clear meaning for all of the parties involved. In reality, however, the perception of space and time is different for every observer. If we think of an object P_A of the category of protocols as the specification of a protocol for a frame of reference A, it is natural to wonder: given a different frame of reference B, how can we obtain a new object P_B that specifies the protocol for it?

The need to adapt the specification to each party is not limited to relativity in physics. For instance, in the example of music, the same score will sound different when played by a piano and a horn. This is because the horn is tuned to the note F, meaning that every note on the score sounds a perfect fifth lower in a horn than it does in a piano, which is tuned to C. Therefore, we need to take this into account when adapting our music specification. In this setting, the question of relativity becomes: given a score for an instrument, how to adapt it to a different tuning system so that it sounds the same? In the music jargon, this is called transposing the score.

Similarly, we wonder: given two objects in the category of protocols, when can we understand them as representing the same protocol but seen from two different frames of reference?

Connection with algebraic quantum field theory

Categories of Hilbert modules, which constitute one of our leading examples, can be seen as naive versions of quantum field theories [42, 26]. In algebraic quantum field theory, systems are pairs $(R, A(R))$, where R is a spacetime region and $A(R)$ is the C^*-algebra of all observables associated to the experiments that potentially could take place in that region [37, Section 2.2.3]. This is somewhat similar to our notion of event: a pair (f, s) of a region s and something f that takes place in it. Perhaps this link could be developed further into a formal connection.

Indefinite causal structure

In quantum theory, and throughout this thesis, it is assumed that there is a predefined, fixed causal structure of spacetime. It has recently been shown, however, that it is in principle possible to physically realise processes that would be incompatible with any fixed causal structure [71, 86], although it is not clear which naturally occurring processes could be of this kind. Roughly speaking, it is theoretically possible to have a quantum superposition of an event A causing event B and B causing A. This has led to an exploration of different ways to weaken the aforementioned underlying assumption on causality, and to study so-called *indefinite* causal structures. Another motivation for this weakening is that, since in general relativity both spacetime and causal structure are in constant dynamic change, conceiving causality as fixed may not be suitable for frameworks for quantum gravity.

We think that the approach we introduce in this thesis has the potential to incorporate ideas about indefinite causal structures. For instance, it may be interesting to let the closure operators, which give the causal cones of regions in our framework, depend on more inputs than just the region. Combined with attempts to account for different observers as mentioned above, and being optimistic, this could be a first step towards a categorical framework for quantum gravity.

Chapter 2

Theory

Subobjects are the generalisation in category theory of the classical notion of algebraic substructures, such as subsets, subrings and vector subspaces. Concretely, a subobject of an object A is an equivalence class of monomorphisms $s\colon S \rightarrowtail A$, where s and s' are identified if they factor through each other.

It is known that subobjects of the tensor unit in the category $\mathbf{Sh}(X)$ of sheaves over a topological space X correspond to the open subsets of X, so they are a categorical way to represent them (see Section 2.2 below). However, in the non-cartesian case of the category $\mathbf{Hilb}_{C_0(X)}$ of Hilbert modules over a topological space X, there are too many subobjects for them to be in a one-to-one correspondence with the open subsets of X. However, if we focus on those subobjects with a certain idempotence property, we can capture exactly the open subsets in this example as well. These subobjects will be called *subunits*, and they are the main focus of this chapter.

The category $\mathbf{Hilb}_{C_0(X)}$ of Hilbert modules (see Definition 2.23) can be characterised in two equivalent ways. Algebraically, it consists of modules over the algebra $C_0(X)$ of complex-valued continuous functions over a base space X (a locally compact Hausdorff topological space). This is roughly the result of replacing, in the definition of Hilbert space, the complex numbers by an arbitrary commutative C^*-algebra, since by Gelfand duality these are all of the form $C_0(X)$ for some X. Geometrically, Hilbert modules are equivalent to bundles of Hilbert spaces over the base space. Each bundle gives a Hilbert space for every base point $x \in X$, which 'varies continuously' with x, and thus we can think of it as a naive version of quantum field theory. In particular, for each $x \in X$ there is a monoidal functor $\mathrm{Loc}_x\colon \mathbf{Hilb}_{C_0(X)} \to \mathbf{Hilb}_{\mathbb{C}}$. For details, see [42].

We work with braided monoidal categories [66], and will sometimes suppress the coherence isomorphisms $\lambda_A\colon I \otimes A \to A$, $\rho_A\colon A \otimes I \to A$, $\sigma_{A,B}\colon A \otimes B \to B \otimes A$ and

$\alpha_{A,B,C} \colon A \otimes (B \otimes C) \to (A \otimes B) \otimes C$, and often abbreviate identity morphisms $1_A \colon A \to A$ simply by A. Whenever we talk about a subobject, we will use a small letter s for a representing monomorphism, and the corresponding capital S for its domain.

This chapter is based on the contents of [26] and [27], both of which were the result of joint work with Chris Heunen and Sean Tull. The former article gives a first short presentation of the ideas, while the latter develops the theory in more depth.

2.1 Subunits

Let us define subunits and prove their main general property: when well-behaved (making the category *firm*), they form a meet-semilattice.

Definition 2.1. A *subunit* in a braided monoidal category $\mathbf{C}$ is a subobject $s \colon S \rightarrowtail I$ of the tensor unit such that $s \otimes S \colon S \otimes S \to I \otimes S$ is an isomorphism[1]. Write $\mathrm{ISub}(\mathbf{C})$ for the collection of subunits in $\mathbf{C}$.

Remark 2.2. Note that if $s \otimes S$ is invertible then so is $S \otimes s$, and that this follows from s being monic, without the need to use the braiding of $\mathbf{C}$.

We could have generalised the previous definition to arbitrary monoidal categories by additionally requiring subunits to be central in the sense that there is a natural isomorphism $(-) \otimes S \Rightarrow S \otimes (-)$. Most results below still hold, but the bureaucracy is not worth the added generality here.

Many results also remain valid when we require $s \otimes S$ not to be invertible but merely split epic, but for simplicity we stick with invertibility.

We begin with some useful observations, mostly adapted from Boyarchenko and Drinfeld [15].

Lemma 2.3. *Let $m \colon A \to B$ and $e \colon B \to A$ satisfy $e \circ m = A$, and $s \colon S \rightarrowtail I$ be a subunit. If $s \otimes B$ is an isomorphism, then so is $s \otimes A$.*

Proof. The diagram below commutes by bifunctoriality of $\otimes$.

$$\begin{tikzcd}
S \otimes A \arrow[r, "S\otimes m"] \arrow[d, "s\otimes A"] & S \otimes B \arrow[r, "S\otimes e"] \arrow[d, "s\otimes B", "\simeq"'] & S \otimes A \arrow[d, "s\otimes A"] \\
I \otimes A \arrow[r, "I\otimes m"] & I \otimes B \arrow[r, "I\otimes e"] & I \otimes A
\end{tikzcd}$$

[1]Boyarchenko and Drinfeld call morphisms $s \colon S \to I$ for which $s \otimes S$ and $S \otimes s$ are isomorphisms *open idempotents* [15], with (the dual of) this notion going back implicitly at least to [54, Exercise 4.2]. In [26] subunits were called *idempotent subunits*.

Both rows compose to the identity, and the middle vertical arrow is an isomorphism. Hence $s \otimes A$ is an isomorphism with inverse $(S \otimes e) \circ (s \otimes B)^{-1} \circ (I \otimes m)$. □

Recall that subobjects of a fixed object always form a partially ordered set, where $s \leq t$ if and only if s factors through t. The following observation characterises this order in another way for subunits.

Lemma 2.4. *A subunit s factors through another t if and only if $S \otimes t$ is invertible, or equivalently, $t \otimes S$ is invertible.*

Proof. Suppose $s = t \circ f$. Set $g = (S \otimes f) \circ (S \otimes s)^{-1} \circ \rho_S^{-1} : S \to S \otimes T$. Then

$$\rho_S \circ (S \otimes t) \circ g = \rho_S \circ (S \otimes s) \circ (S \otimes s)^{-1} \circ {\rho_S}^{-1} = S.$$

Idempotence of t makes $S \otimes T \otimes t : S \otimes T \otimes T \to S \otimes T \otimes I$ an isomorphism. Hence, by the right-handed version of Lemma 2.3, so is $S \otimes t$. A symmetric argument makes $t \otimes S$ invertible.

Conversely, suppose $S \otimes t$ is an isomorphism. Because the diagram

$$\begin{tikzcd}
S \otimes T \arrow[r, "s \otimes T"] \arrow[d, "S \otimes t"] & I \otimes T \arrow[r, "\rho_T"] \arrow[d, "I \otimes t"] & T \arrow[d, "t"] \\
S \otimes I \arrow[r, "s \otimes I"] & I \otimes I \arrow[r, "\rho_I"] & I
\end{tikzcd}$$

commutes, the bottom row $s \circ \rho_S$ factors through the right vertical arrow t, whence so does s. □

It follows from Lemma 2.4 that subunits are determined by their domain: if $s, s' : S \rightarrowtail I$ are subunits, then $s' = s \circ f$ for a unique f, which is an isomorphism. This justifies our convention to use the same letter for a subunit and its domain.

We want the tensor product of subunits to result in a new subunit, but in general it is not guaranteed that the product of two monomorphisms is again a monomorphism. So that we are able to do this, and for the resulting theory to work smoothly, we impose the condition of *firmness* on the categories we deal with.

Definition 2.5. A category is called *firm* when it is braided monoidal and $s \otimes T : S \otimes T \to I \otimes T$ is a monomorphism whenever s and t are subunits.

Remark 2.6. The name *firm* is chosen after Quillen [76], who employs it as a natural condition for nonunital rings to make up for a missing unit. The previous definition extends the term to the category of nonunital rings; see Example 2.19 below. Note, however, that a firm category has genuine identity morphisms and a genuine tensor unit. Firmness is a very mild condition: Example 2.72 below gives a category that is not firm, but we know of no other 'naturally occurring' categories that are not firm.

Lemma 2.7. *Any co-closed braided monoidal category is firm.*

Proof. Each functor $(-) \otimes T$ is a right adjoint and so preserves limits and hence monomorphisms. Hence whenever s is monic so is $s \otimes T$. $\square$

In particular, a $*$-autonomous category is firm, as is a compact category.

Remark 2.8. In the following, we will completely disregard size issues, and pretend $\mathrm{ISub}(\mathbf{C})$ is a set, as in our main examples.

Proposition 2.9. *The subunits in a firm category form a semilattice, with largest element I, meets given by*

$$\big(s \colon S \rightarrowtail I\big) \wedge \big(t \colon T \rightarrowtail I\big) = \big(\lambda_I \circ (s \otimes t) \colon S \otimes T \rightarrowtail I\big),$$

and the usual order of subobjects.

Proof. First observe that $s \otimes t = (I \otimes t) \circ (s \otimes T)$ is monic, because $I \otimes t = \lambda_I^{-1} \circ t \circ \lambda_T$ is monic, and $s \otimes T$ is monic by firmness. It is easily seen to be idempotent using the braiding, and hence it is a well-defined subunit.

Next, we show that $\mathrm{ISub}(\mathbf{C})$ is an idempotent commutative monoid under $\wedge$ and I. The subunit I is a unit as $I \otimes s = \lambda_I \circ (I \otimes s) = s \circ \lambda_S$ represents the same subobject as s, and similarly $I \otimes s$ represents the same subobject as s because $\rho_I = \lambda_I$. An analogous argument using coherence establishes associativity. For commutativity, use the braiding to observe that $s \otimes t$ and $t \otimes s$ represent the same subobject. For idempotence note that $s \otimes s$ and s represent the same subobject because $\lambda_I \circ (s \otimes s) = s \circ \rho_S \circ (S \otimes s)$.

Hence $\mathrm{ISub}(\mathbf{C})$ is a semilattice where s is below t if and only if $s = s \wedge t$. Finally, we show that this order is the same as the usual order of subobjects. On the one hand, if s and $s \otimes t$ represent the same subobject, then $S \simeq S \otimes T$, making $S \otimes t$ an isomorphism and so $s \leq t$ by Lemma 2.4.

$$\begin{tikzcd} T \arrow[dr, tail, "t"] & \\ & I \\ S \arrow[uu, tail, dashed] \arrow[ur, tail, "s"'] & \end{tikzcd} \qquad \Longleftrightarrow \qquad \begin{tikzcd} S \otimes T \arrow[r, tail, "s \otimes t"] & I \otimes I \arrow[d, "\simeq"', "\lambda_I"] \\ S \arrow[u, tail, "\simeq"] \arrow[r, tail, "s"'] & I \end{tikzcd}$$

On the other hand, if $s \leq t$ then by the same lemma $S \otimes t$ is an isomorphism with $s = \lambda_I \circ (s \otimes t) \circ (S \otimes t)^{-1} \otimes \rho_S^{-1}$, and so both subobjects are equal. $\square$

2.2 Examples

This section determines the subunits of four families of examples: cartesian categories, like sheaves over a topological space; commutative unital quantales; firm modules over a nonunital ring; and Hilbert modules over a nonunital commutative C*-algebra.

Cartesian categories

We start with examples in which the tensor product is in fact a product.

Example 2.10. Any cartesian category $\mathbf{C}$ is firm, and $\mathrm{ISub}(\mathbf{C})$ consists of the subobjects of the terminal object.

In particular, if X is a topological space, then subunits in its category of sheaves $\mathrm{Sh}(X)$ correspond to open subsets of X [13, Corollary 2.2.16].

Proof. Let $s\colon S \rightarrowtail 1$ be a subterminal object. Let $\Delta = \langle S, S\rangle\colon S \to S \times S$ be the diagonal and write $\pi_i\colon A_1 \times A_2 \to A_i$ for the projections. Then $(s \times S) \circ \Delta \circ \pi_2 = \pi_2^{-1} \circ S \circ \pi_2 = 1 \times S$. Now, the unique map s of type $S \to 1$ is monic precisely when any two parallel morphisms into S are equal. Hence $\pi_i \circ \Delta \circ \pi_2 \circ (s \times S) = \pi_i$, and so $\Delta \circ \pi_2 \circ (s \times S) = \langle \pi_1, \pi_2\rangle = S \times S$. Thus $s \times S$ is automatically invertible.

Finally, suppose $s_i\colon S_i \rightarrowtail 1$ for $i = 1,2$ are monic, and that $f, g\colon A \to S_1 \times S_2$ satisfy $(s_1 \times s_2) \circ f = (s_1 \times s_2) \circ g$. Postcomposing with π_i shows that $s_i \circ \pi_i \circ f = s_i \circ \pi_i \circ g$, whence $\pi_i \circ f = \pi_i \circ g$ and so $f = g$. This establishes firmness. □

Semilattices

Next we consider examples that are degenerate in another sense: firm categories in which there is at most one morphism between two given objects.

Example 2.11. Any semilattice $(L, \wedge, 1)$ forms a strict symmetric monoidal category: objects are $x \in L$, there is a unique morphism $x \to y$ if $x \leq y$, tensor product is given by meet, and tensor unit is $I = 1$. Every morphism is monic so this monoidal category is firm, and its semilattice of (idempotent) subunits is $(L, \wedge, 1)$.

This gives the free firm category on a semilattice. More precisely, this construction is left adjoint to the functor from the category **Firm** of firm categories with (strong) monoidal subunit-preserving functors to the category **SLat** of semilattices and their homomorphisms, which takes subunits.

$$\begin{tikzcd} \mathbf{SLat} \arrow[r, shift left] \arrow[r, phantom, "\perp"] & \mathbf{Firm} \arrow[l, shift left, "\mathrm{ISub}"] \end{tikzcd}$$

Quantales

We move on to more interesting examples, namely special kinds of semilattices like frames and quantales.

Definition 2.12. A *frame* is a complete lattice in which finite joins distribute over suprema. A morphism of frames is a function that preserves $\bigvee$, $\wedge$, and 1. Frames and their morphisms form a category **Frame**.

The prototypical example of a frame is the collection of open sets of a topological space [48]. Frames may be generalised as follows [78].

Definition 2.13. A *quantale* is a monoid in the category of complete lattices. More precisely, it is a partially ordered set Q that has all suprema, that has a multiplication $Q \times Q \to Q$, and that has an element e, such that:

$$a\big(\bigvee b_i\big) = \bigvee ab_i, \qquad \big(\bigvee a_i\big)b = \bigvee a_i b, \qquad ae = a = ea.$$

A morphism of quantales is a function that preserves $\bigvee$, $\cdot$, and e. A quantale is *commutative* when $ab = ba$ for all $a, b \in Q$. Commutative quantales and their morphisms for a category **cQuant**.

Equivalently, a frame is a commutative quantale in which the multiplication is idempotent.

Any quantale may be regarded as a monoidal category, whose objects are elements of the quantale, where the (composition of) morphisms is induced by the partial order, and the tensor product is induced by the multiplication. This monoidal category is firm, but only braided if the quantale is commutative.

Example 2.14. Taking subunits is right adjoint to the inclusion:

$$\begin{tikzcd} \mathbf{Frame} \arrow[r, shift left] \arrow[r, phantom, "\perp"] & \mathbf{cQuant} \arrow[l, shift left, "\mathrm{ISub}"] \\ \{q \in Q \mid q^2 = q \leq e\} & Q \arrow[l, mapsto] \end{tikzcd}$$

Proof. We first prove that $\mathrm{ISub}(Q)$ is a well-defined frame. If $q_i \in \mathrm{ISub}(Q)$,

$$(\bigvee q_i)^2 = \bigvee_{i,j} q_i q_j \le \bigvee_{i,j} q_i e = \bigvee_i q_i = \bigvee_i q_i q_i \le \bigvee_{i,j} q_i q_j = (\bigvee q_i)^2$$

and $\bigvee q_i \le \bigvee_i e = e$, so $\bigvee q_i \in \mathrm{ISub}(Q)$. Moreover, if $p, q \in \mathrm{ISub}(Q)$, then pq is again below e and is idempotent by commutativity of Q. Moreover $pq = p \wedge q$ in $\mathrm{ISub}(Q)$: if $o \in \mathrm{ISub}(Q)$ has $o \le pq$ then $o \le pq \le pe = p$ and similarly $o \le q$; and conversely if $o \le p$ and $o \le q$ then $o = oo \le pq$. Since quantale multiplication distributes over suprema, then so do finite meets.

For the adjunction, observe that if F is a frame and Q is a commutative quantale, then $F = \mathrm{ISub}(F)$ and any morphism $F \longrightarrow Q$ of quantales restricts to a unique morphism of frames $F \longrightarrow \mathrm{ISub}(Q)$. □

Remark 2.15. Examples 2.10 and 2.14 show that subunits do not capture all possible topological content in the following, more conventional contexts.

For a Grothendieck topos, subunits form the poset of internal truth values, which does not suffice to reconstruct the category, which may itself be said to embody a notion of topological space.

For the commutative quantale, $[0, \infty]$ under multiplication and the usual order, the subunits form the two-element Boolean algebra, which is clearly far poorer than the quantale itself.

Example 2.16. If M is a monoid, then its (right) ideals form a unital quantale Q with multiplication $IJ = \{xy \mid x \in I, y \in J\}$ and unit M itself. When M is commutative, so is Q, and $\mathrm{ISub}(Q)$ consists of all ideals satisfying $I = II$.

Example 2.17. If R is a commutative ring, then its additive subgroups form a unital commutative quantale Q with multiplication $GH = \{x_1 y_1 + \cdots + x_n y_n \mid x_i \in G, y_i \in H\}$, supremum $\bigvee G_i = \{\sum_{j \in J} x_j \mid x_j \in G_j \text{ for } J \subseteq I \text{ finite}\}$, and unit

$$\mathbb{Z}1 = \{0, 1, -1, 1+1, -1-1, 1+1+1, -1-1-1, \ldots\}$$

Then $G \le H$ iff $G \subseteq H$ and $\mathrm{ISub}(Q)$ consists of those subgroups G such that $G \subseteq G \cdot G$ and $G \subseteq \mathbb{Z}1$. The latter means that G must be of the form $n\mathbb{Z}1$ for some $n \in \mathbb{N}$. The former then means that $n1 = n^2 y1$ for some $y \in \mathbb{Z}$. Thus

$$\mathrm{ISub}(Q) = \{n\mathbb{Z}1 \mid n \in \mathbb{N}, \exists y \in \mathbb{Z} \colon n1 = n^2 y1\}$$

Modules

Another example of a monoidal category is that of modules over a ring. We have to take some pains to treat nonunital rings.

Definition 2.18. A commutative ring R is *firm* when its multiplication is a bijection $R \otimes_R R \to R$, and *nondegenerate* when $r \in R$ vanishes as soon as $rs = 0$ for all $s \in R$. Any unital ring is firm and nondegenerate, but examples also include infinite direct sums $\bigoplus_{n\in\mathbb{N}} R_n$ of unital rings R_n. Firm rings R are *idempotent*: they equal $R^2 = \{\sum_{i=1}^n r_i' r_i'' \mid r_i', r_i'' \in R\}$. Let R be a nondegenerate firm commutative ring. An R-module E is *firm* when the scalar multiplication is a bijection $E \otimes_R R \to E$ [76], and *nondegenerate* when $x \in E$ vanishes as soon as $xr = 0$ for all $r \in R$. If R is unital, then every R-module is firm and nondegenerate. Nondegenerate firm R-modules and linear maps form a monoidal category $\mathbf{FMod}_R$.

Example 2.19. The subunits in $\mathbf{FMod}_R$ correspond to *nondegenerate firm idempotent ideals*: ideals $S \subseteq R$ that are idempotent as rings, and nondegenerate and firm as R-modules. Any ideal that is unital as a ring is a nondegenerate firm idempotent ideal. The category $\mathbf{FMod}_R$ is firm.

Proof. Monomorphisms are injective by nondegeneracy, so every subunit is a nondegenerate firm R-submodule of R, that is, a nondegenerate firm ideal. Because the inclusion $S \otimes S \to R \otimes S$ is surjective and S is firm, the map $S \otimes S \to S$ given by $s' \otimes s'' \mapsto s's''$ is surjective. Thus S is idempotent.

Conversely, let S be a nondegenerate firm idempotent ideal of R. The inclusion $S \otimes S \to R \otimes S$ is surjective, as $r \otimes s \in R \otimes S$ can be written as $r \otimes s's'' = rs' \otimes s'' \in S \otimes S$. Hence S is a subunit.

Next suppose ideal S is unital (with generally $1_S \neq 1_R$ if R is unital). Then $S \otimes R \to S$ given by $s \otimes r \mapsto sr$ is bijective: surjective as $1_S \otimes s \mapsto 1_S s = s$; and injective as $s \otimes r = 1_S \otimes sr = 1_S \otimes 0 = 0$ if $sr = 0$. Hence S is firm and nondegenerate. Any $s \in S$ can be written as $s = s1_S \in S^2$, so S is idempotent.

Finally, to see that the category is firm, let $S, T \subseteq R$ be nondegenerate firm idempotent ideals. We need to show that the map $S \otimes T \mapsto R \otimes T$ given by $s \otimes t \mapsto s \otimes t$ is injective. Because T is firm, it suffices that multiplication $S \otimes T \to S$ given by $s \otimes t \mapsto st$ is injective, which holds because S is firm. □

The previous example generalises to commutative nonunital bialgebras in any symmetric monoidal category.

Example 2.20. Let **C** be a symmetric monoidal category. A *commutative nonunital bialgebra* in **C** is an object M together with an associative multiplication $\mu\colon M\otimes M \to M$ and a comonoid $\delta\colon M \to M\otimes M$, $\varepsilon\colon M \to I$, for which μ and δ are commutative and satisfy both $\varepsilon\circ\mu = \varepsilon\otimes\varepsilon$ and the bialgebra law:

$$(\mu\otimes\mu)\circ(M\otimes\sigma\otimes M)\circ(\delta\otimes\delta) = \delta\circ\mu$$

We define a braided monoidal category $\mathbf{Mod}_M$ where objects are $\alpha\colon M\otimes A \to A$ satisfying $\alpha\circ(\mu\otimes A) = \alpha\circ(M\otimes\alpha)$, with morphisms and $\otimes$ all defined as for modules over a (unital) commutative bialgebra (see e.g. [38, 2.2,2.3]). The category $\mathbf{Mod}_M$ is firm when **C** is, and its subunits correspond to *firm ideals*: monomorphisms $s\colon S \rightarrowtail M$ such that

$$\begin{tikzcd} M\otimes S \arrow[r, "M\otimes s"] \arrow[d, dashed] & M\otimes M \arrow[d, "\mu"] \\ S \arrow[r, tail, "s"'] & M \end{tikzcd}$$

and $\varepsilon\otimes S$ and $s\otimes S$ are isomorphisms.

We next instantiate the previous example in two special cases: in the monoidal categories of semilattices and of quantales.

Example 2.21. Any semilattice M is a nondegenerate nonunital bialgebra in **SLat**. In $\mathbf{Mod}_M$ objects are semilattices A with functions $\alpha\colon M\times A \to A$ which respect $\wedge$ in each argument and satisfy $\alpha(x\wedge y, a) = \alpha(x, \alpha(y,a))$. Subobjects of the tensor unit correspond to subsets $S\subseteq M$ which are ideals under $\wedge$, or equivalently downward-closed. Because $x\otimes y = (x\wedge x)\otimes y = x\otimes(x\wedge y) \in S\otimes S$, we have $S\otimes S = S\otimes M$, and every subobject of the tensor unit is a subunit.

Example 2.22. Any commutative unital quantale M is a nondegenerate nonunital bialgebra in the category of complete lattices. $\mathbf{Mod}_M$ then consists of complete lattices A with functions $\alpha\colon M\times A \to A$ preserving arbitrary suprema in each argument and with $\alpha(x, \alpha(y,a)) = \alpha(xy, a)$. Subobjects of the tensor unit are subsets $S\subseteq M$ closed under both arbitrary suprema and multiplication with elements of M. Subunits furthermore have that for every $r\in S$ and $x\in M$ there exist $s_i, t_i \in S$ with $r\otimes x = \bigvee s_i\otimes t_i$. For example, if $M = [0,\infty]$ under addition with the opposite ordering, subunits include $\emptyset, \{\infty\}$, $\{0,\infty\}$, $(0,\infty]$, and $[0,\infty]$.

Hilbert modules

The above examples of module categories were all algebraic in nature. Our next suite of examples is more analytic.

Definition 2.23. Fix a locally compact Hausdorff space X. It induces a commutative C*-algebra

$$C_0(X) = \{f\colon X \to \mathbb{C} \text{ continuous} \mid \forall \varepsilon > 0\, \exists K \subseteq X \text{ compact}\colon |f(X \setminus K)| < \varepsilon\}.$$

A *Hilbert module* is a $C_0(X)$-module A with a map $\langle - \mid - \rangle\colon A \times A \to C_0(X)$ that is $C_0(X)$-linear in the second variable, satisfies $\langle a \mid b\rangle = \langle b \mid a\rangle^*$, and $\langle a \mid a\rangle \geq 0$ with equality only if $a = 0$, and makes A complete in the norm $\|a\|_A^2 = \sup_{x\in X}\langle a \mid a\rangle(x)$. A function $f\colon A \to B$ between Hilbert $C_0(X)$-modules is *bounded* when $\|f(a)\|_F \leq \|f\|\|a\|_A$ for some $\|f\| \in \mathbb{R}$. Here we will focus on *contractions*, i.e. those bounded functions with $\|f\| \leq 1$.

Hilbert modules were first introduced by Kaplansky [52] and studied by many others, including Rieffel [77], Kasparov [55], and Mulvey [70]. For more information we refer to [62].

The category $\mathbf{Hilb}_{C_0(X)}$ of Hilbert $C_0(X)$-modules and contractive $C_0(X)$-linear maps is not abelian, not complete, and not cocomplete [40]. Nevertheless, $\mathbf{Hilb}_{C_0(X)}$ is symmetric monoidal [42, Proposition 2.2]. Here $A \otimes B$ is constructed as follows: consider the algebraic tensor product of $C_0(X)$-modules, and complete it to a Hilbert module with inner product $\langle a \otimes b \mid a' \otimes b'\rangle$ given by $\langle a \mid a'\rangle\langle b \mid b'\rangle$. The tensor unit is $C_0(X)$ itself, which forms a Hilbert $C_0(X)$-module under the inner product $\langle f \mid g\rangle(x) = f(x)^*g(x)$.

Example 2.24. $\mathbf{Hilb}_{C_0(X)}$ is firm, and its subunits are

$$\{f \in C_0(X) \mid f(X \setminus U) = 0\} \simeq C_0(U) \tag{2.1}$$

for open subsets $U \subseteq X$.

Proof. If U is an open subset of X, we may indeed identify $C_0(U)$ with the closed ideal of $C_0(X)$ in (2.1): if $f \in C_0(U)$, then its extension by zero on $X \setminus U$ is in $C_0(X)$, and conversely, if $f \in C_0(X)$ is zero outside U, then its restriction to U is in $C_0(U)$. Moreover, note that the canonical map $C_0(X) \otimes C_0(X) \to C_0(X)$ is always an isomorphism as $C_0(X)$ is the tensor unit, and hence the same holds for $C_0(U)$. Thus $C_0(U)$ is a subunit in $\mathbf{Hilb}_{C_0(X)}$.

For the converse, let $s\colon S \rightarrowtail C_0(X)$ be a subunit in $\mathbf{Hilb}_{C_0(X)}$. We will show that $s(S)$ is a closed ideal in $C_0(X)$, and therefore of the form $C_0(U)$ for some open subset $U \subseteq X$. It is an ideal because s is $C_0(X)$-linear. To see that it is closed, let $g \in s(S)$. Then

$$\begin{aligned}
\|g\|_S^4 &= \|\langle g \mid g\rangle_S^2\|_{C_0(X)} = \|\langle g \mid g\rangle_S\langle g \mid g\rangle_S\|_{C_0(X)} \\
&= \|\langle g\otimes g \mid g\otimes g\rangle_{C_0(X)}\|_{C_0(X)} = \|g\otimes g\|_S^2 \\
&\leq \|\rho_S^{-1}\|^2\|g^2\|_S = \|\rho_S^{-1}\|^2\|\langle g \mid g\rangle_S g^* g\|_{C_0(X)} \\
&\leq \|\rho_S^{-1}\|^2\|g\|_S^2\|g\|_{C_0(X)}^2
\end{aligned}$$

and therefore $\|g\|_S \leq \|\rho_S^{-1}\|^2\|g\|_{C_0(X)}^2$. Because s is bounded, it is thus an equivalence of normed spaces between $(S, \|-\|_S)$ and $(s(S), \|-\|_{C_0(X)})$. Since the former is complete, so is the latter. Firmness follows from Example 2.35 later. □

The category $\mathbf{Hilb}_{C_0(X)}$ can be adapted to form a dagger category by considering (not necessarily contractive) bounded maps between Hilbert modules that are *adjointable*. In that case only clopen subsets of X correspond to subunits [42, Lemma 3.3].

Not every subobject of the tensor unit in $\mathbf{Hilb}_{C_0(X)}$ is induced by an open subset $U \subseteq X$, and so the condition of Definition 2.1 is not redundant.

Example 2.25. Let $X = [0,1]$. If $f \in C_0(X)$, write $\hat{f} \in C_0(X)$ for the map $x \mapsto xf(x)$. Then $S = \{\hat{f} \mid f \in A\}$ is a subobject of $A = C_0(X)$ in $\mathbf{Hilb}_{C_0(X)}$ under $\langle \hat{f} \mid \hat{g}\rangle_S = \langle f \mid g\rangle_A$, that is not closed under $\|-\|_A$.

Proof. Clearly S is a $C_0(X)$-module, and $\langle - \mid -\rangle_S$ is sesquilinear. Moreover S is complete: $\hat{f}_n$ is a Cauchy sequence in S if and only if f_n is a Cauchy sequence in A, in which case it converges in A to some f, and so $\hat{f}_n$ converges to $\hat{f}$ in S. Thus S is a well-defined Hilbert module. The inclusion $S \hookrightarrow A$ is bounded and injective, and hence a well-defined monomorphism. In fact, A is a C*-algebra, and S is an ideal. The closure of S in A is the closed ideal $\{f \in C_0(X) \mid f(0) = 0\}$, corresponding to the closed subset $\{0\} \subseteq X$. It contains the function $x \mapsto \sqrt{x}$ while S does not, and so S is not closed. □

2.3 Restriction

Regarding subunits as open subsets of an (imagined) base space, the idea of restriction to such an open subset makes sense. For example, if U is an open subset of a locally compact Hausdorff space X, then any $C_0(X)$-module E induces, by restriction of scalars, a $C_0(U)$-module unitarily equivalent to $E \otimes C_0(U)$ (see [26, Lemma 32]), and any sheaf over X induces a sheaf over U. More generally, any subunit in a topos induces an open subtopos. This section shows that this restriction behaves well in any monoidal category.

Definition 2.26. A morphism $f\colon A \to B$ *restricts to* a subunit $s\colon S \to I$ when it factors through $\lambda_B \circ (s \otimes B)$.

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, dashed] & B \\ S \otimes B \arrow[r, "s \otimes B"'] & I \otimes B \arrow[u, "\lambda_B"'] \end{tikzcd}$$

As a special case, we can consider to which subunits identity morphisms restrict [16, Lemma 1.3].

Proposition 2.27. *The following are equivalent for an object A and subunit s:*

(a) $s \otimes A\colon S \otimes A \to I \otimes A$ is an isomorphism;

(b) there is an isomorphism $S \otimes A \simeq A$;

(c) there is an isomorphism $S \otimes B \simeq A$ for some object B;

(d) the identity $A \to A$ restricts to s.

Proof. Trivially (a) $\implies$ (b) $\implies$ (c). For (c) $\implies$ (d): because s is a subunit, $s \otimes S \otimes A$ is an isomorphism, so if $S \otimes B \simeq A$ then also $s \otimes A$ is an isomorphism by Lemma 2.3. For (d) $\implies$ (a): if A factors through $s \otimes A$, then because s is a subunit $s \otimes S \otimes A$ is an isomorphism, and hence so is $s \otimes A$ by Lemma 2.3. □

The following observation is simple, but effective in applications [26].

Lemma 2.28. *Let $s\colon S \to I$ and $t\colon T \to I$ be subunits in a firm category. If f restricts to s, and g restricts to t, then $f \circ g$ and $f \otimes g$ restrict to $s \wedge t$.*

Proof. Straightforward. □

In particular, if A or B restrict to a subunit s, then so does any map $A \to B$. It also follows that restriction respects retractions: if $e \circ m = 1$, then m restricts to s if and only if e does.

Definition 2.29. Let s be a subunit in a monoidal category $\mathbf{C}$. Define the *restriction* of $\mathbf{C}$ to s, denoted by $\mathbf{C}|_s$, to be the full subcategory of $\mathbf{C}$ of objects A for which $s \otimes A$ is an isomorphism.

Proposition 2.30. *If s is a subunit in a monoidal category $\mathbf{C}$, then $\mathbf{C}|_s$ is a coreflective monoidal subcategory of $\mathbf{C}$.*

$$\begin{tikzcd} \mathbf{C} \arrow[r, shift left] \arrow[r, phantom, "\top"] & \mathbf{C}|_s \arrow[l, hook, shift left] \end{tikzcd}$$

The right adjoint $\mathbf{C} \to \mathbf{C}|_s$, given by $A \mapsto S \otimes A$ and $f \mapsto S \otimes f$, is also called restriction *to s.*

Proof. First, if $A \in \mathbf{C}$, note that $S \otimes A$ is indeed in $\mathbf{C}|_s$ because $s \otimes S \otimes A$ is an isomorphism as s is a subunit. Similarly, $\mathbf{C}|_s$ is a monoidal subcategory of $\mathbf{C}$. Finally, there is a natural bijection

$$\begin{aligned} \mathbf{C}(A, B) &\simeq \mathbf{C}|_s(A, S \otimes B) \\ f &\mapsto (s \otimes f) \circ (s \otimes A)^{-1} \circ \rho_A^{-1} \\ \lambda_B \circ (s \otimes B) \circ g &\leftarrow\!\shortmid g \end{aligned}$$

for $A \in \mathbf{C}|_s$ and $B \in \mathbf{C}$. So restriction is right adjoint to inclusion. For monoidality, see [47, Theorem 5]; both functors are (strong) monoidal when $\mathbf{C}|_s$ has tensor unit S and tensor product inherited from $\mathbf{C}$. □

Remark 2.31. The previous result motivates our terminology; a subunit s in $\mathbf{C}$ is precisely a subobject of I with the property that it may form the tensor unit of a monoidal subcategory of $\mathbf{C}$, namely $\mathbf{C}|_s$.

Example 2.32. Let L be a semilattice, regarded as a firm category as in Example 2.11. For a subset $U \subseteq L$ we define $\downarrow U = \{x \in L \mid x \leq u \text{ for some } u \in U\}$. Then for $s \in L$, the restriction $\mathbf{C}|_s$ is the subsemilattice $\downarrow s = \downarrow\{s\}$.

Example 2.33. Let L be a frame. A subunit in $\mathrm{Sh}(L)$ is just an element $s \in L$, and a morphism $f \colon A \Rightarrow B$ restricts to it precisely when $A(x) = \emptyset$ for $x \nleq s$.

Example 2.34. Let S be a nondegenerate firm idempotent ideal of a nondegenerate firm commutative ring R. Then $\mathbf{FMod}_R|_S$ is monoidally equivalent to $\mathbf{FMod}_S$.

Proof. Send A in $\mathbf{FMod}_R|_S$ to A with S-module structure $a \cdot s := as$, and send an R-linear map f to f. This defines a functor $\mathbf{FMod}_R|_S \longrightarrow \mathbf{FMod}_S$. In the other direction, a firm S-module $B \simeq B \otimes_S S$ has firm R-module structure $(b \otimes s) \cdot r := b \otimes (sr)$ because S is idempotent, and if g is an S-linear map then $g \otimes_S S$ is R-linear. This defines a functor $\mathbf{FMod}_S \longrightarrow \mathbf{FMod}_R|_S$. Composing both functors sends a firm R-module A to $A \otimes_S S \simeq A \otimes_R R \simeq A$, and a firm S-module B to $B \otimes_S S \simeq B$. □

Example 2.35. For any Hilbert $C_0(X)$-module A and subunit $C_0(U)$ induced by an open subset $U \subseteq X$, the module $A \otimes C_0(U)$ is isomorphic to its submodule

$$A|_U = \{a \in A \mid \langle a \mid a \rangle \in C_0(U)\}$$

again viewing $C_0(U)$ as a closed ideal of $C_0(X)$ via (2.1). Hence in $\mathbf{Hilb}_{C_0(X)}$ a morphism $f \colon A \longrightarrow B$ restricts to this subunit when $\langle f(a) \mid f(a) \rangle \in C_0(U)$ for all $a \in A$.

Restricting $\mathbf{Hilb}_{C_0(X)}$ to this subunit thus gives the full subcategory of modules A with $A = A|_U$. This is nearly, but not quite, $\mathbf{Hilb}_{C_0(U)}$: any such module also forms a $C_0(U)$-module, but conversely there is no obvious way to extend the action of scalars on a general $C_0(U)$-module to make it a $C_0(X)$-module. There is a so-called *local* adjunction between $\mathbf{Hilb}_{C_0(X)}|_{C_0(U)}$ and $\mathbf{Hilb}_{C_0(U)}$, which is only an adjunction when U is clopen [19, Proposition 4.3].

Proof. Write $S = C_0(U)$. We first prove that $A \in \mathbf{Hilb}_{C_0(X)}|_S$ if and only if $|a| \in C_0(U)$ for all $a \in A$, where $|a|^2 = \langle a, a \rangle$. On the one hand, if $a \in A$ and $f \in S$ then $|a \otimes f|(X \setminus U) = |a||f|(X \setminus U) = 0$. Therefore $|a| \in C_0(U)$ for all $a \in A \otimes S$. Because $A \otimes S \simeq A$ is invertible, $|a| \in C_0(U)$ for all $a \in A$.

On the other hand, suppose that $|a| \in C_0(U) = 0$ for all $a \in A$. We are to show that the morphism $A \otimes S \longrightarrow A$ given by $a \otimes f \mapsto af$ is bijective. To see injectivity, let $f \in S$ and $a \in A$, and suppose that $af = 0$. Then $|a| \cdot |f| = |af| = 0$, so for all $x \in U$ either $|a|(x) = 0$ or $f(x) = 0$. So $|a \otimes f|(U) = 0$, and hence $a \otimes f = 0$. To see surjectivity, let $a \in A$. Then $|a|(x) = 0$ for all $x \in X \setminus U$. So $a = \lim a f_n$ for an approximate unit f_n of S. But that means a is the image of $\lim a \otimes f_n$. □

Above we restricted along one individual subunit s. Next we investigate the structure of the family of these functors when s varies.

Definition 2.36. [32] Let $\mathbf{C}$ be a category and $(\mathbf{E}, \otimes, 1)$ a monoidal category. Denote by $[\mathbf{C}, \mathbf{C}]$ the monoidal category of endofunctors of $\mathbf{C}$ with $F \otimes G = G \circ F$. An $\mathbf{E}$-*graded monad* on $\mathbf{C}$ is a lax monoidal functor $T \colon \mathbf{E} \to [\mathbf{C}, \mathbf{C}]$. More concretely, an $\mathbf{E}$-graded monad consists of:

- a functor $T \colon \mathbf{E} \to [\mathbf{C}, \mathbf{C}]$;
- a natural transformation $\eta \colon 1_{\mathbf{C}} \Rightarrow T(1)$;
- a natural transformation $\mu_{s,t} \colon T(t) \circ T(s) \to T(s \otimes t)$ for all s, t in $\mathbf{E}$;

making the following diagrams commute for all r, s, t in $\mathbf{E}$.

$$\begin{tikzcd}
& T(t) \circ T(s) \circ T(r) \arrow[dl, "\mu_{r,s} \otimes 1_{T(t)}"'] \arrow[dr, "1_{T(r)} \otimes \mu_{s,t}"] & \\
T(t) \circ T(r \otimes s) \arrow[d, "\mu_{r \otimes s, t}"'] & & T(t \otimes s) \circ T(r) \arrow[d, "\mu_{r, s \otimes t}"] \\
T((r \otimes s) \otimes t) \arrow[rr, "T(\alpha_{r,s,t})"] & & T(r \otimes (s \otimes t))
\end{tikzcd}$$

$$\begin{tikzcd}
T(s) \circ 1_{\mathbf{C}} \arrow[r, "\eta \otimes 1_{T(s)}"'] \arrow[d, equal] & T(s) \circ T(1) \arrow[d, "\mu_{1,s}"] \\
T(s) & T(1 \otimes s) \arrow[l, "T(\lambda_s)"]
\end{tikzcd}$$

$$\begin{tikzcd}
1_{\mathbf{C}} \circ T(s) \arrow[r, "1_{T(s)} \otimes \eta"'] \arrow[d, equal] & T(1) \circ T(s) \arrow[d, "\mu_{s,1}"] \\
T(s) & T(s \otimes 1) \arrow[l, "T(\rho_s)"]
\end{tikzcd}$$

Theorem 2.37. *Let $\mathbf{C}$ be a monoidal category. Restriction is a monad graded over the subunits, when we do not identify monomorphisms representing the same subunit. More precisely, it is an $\mathbf{E}$-graded monad, where $\mathbf{E}$ has as objects monomorphisms $s \colon S \rightarrowtail I$ in $\mathbf{C}$ with $s \otimes S$ an isomorphism, and as morphisms $f \colon s \to t$ those f in $\mathbf{C}$ with $s = t \circ f$.*

Proof. The functor $\mathbf{E} \to [\mathbf{C}, \mathbf{C}]$ sends $s \colon S \rightarrowtail I$ to $(-) \otimes S$, and f to the natural transformation $1_{(-)} \otimes f$. The natural transformation $\eta_E \colon E \to E \otimes I$ is given by ρ_E^{-1}. The family of natural transformations $\mu_{s,t} \colon ((-) \otimes S) \otimes T \to (-) \otimes (S \otimes T)$ is given by $\alpha_{(-),S,T}$. Associativity and unitality diagrams follow. □

We end this section by giving two characterisations of subunits in terms that are perhaps more well-known. The first characterisation is in terms of idempotent comonads.

Definition 2.38. A *restriction comonad* on a monoidal category $\mathbf{C}$ is a monoidal comonad $F\colon \mathbf{C} \to \mathbf{C}$:

- whose comultiplication $\delta\colon F \Rightarrow F^2$ is invertible;
- whose counit $\varepsilon\colon F \to 1_{\mathbf{C}}$ has a monic unit component $\varepsilon_I\colon F(I) \rightarrowtail I$.

Proposition 2.39. *Let $\mathbf{C}$ be a braided monoidal category. There is a bijection between subunits in $\mathbf{C}$ and restriction comonads on $\mathbf{C}$.*

Proof. If $s\colon S \rightarrowtail I$ is a subunit, then $F(A) = S \otimes A$ defines a comonad by Proposition 2.30. Its comultiplication is given by $\delta_A = (\lambda_{S\otimes A} \circ (s \otimes S \otimes A))^{-1}$, by definition being an isomorphism. Its counit is given by $\varepsilon_A = \lambda_A \circ (s \otimes A)$. Because $\rho_I = \lambda_I$, its component $\varepsilon_I = \lambda_I \circ (s \otimes I) = \rho_I \circ (s \otimes I) = s \circ \rho_S$ is monic.

Conversely, if F is a restriction monad, then $\varepsilon_I\colon F(I) \rightarrowtail I$ is a subobject of the tensor unit. Writing $\varphi_{A,B}\colon A \otimes F(B) \to F(A \otimes B)$ for the coherence maps, and $\psi_{A,B} = F(\sigma) \circ \varphi_{B,A} \circ \sigma\colon F(A) \otimes B \to F(A \otimes B)$ for its induced symmetric version, the insides of the following diagram commute:

$$\begin{tikzcd}
F(I)\otimes F(I) \arrow[rr, "F(I)\otimes\varepsilon_I"] \arrow[d, "\varphi_{F(I),I}"'] & & F(I)\otimes I \\
F(F(I)\otimes I) \arrow[r, equal] \arrow[ddd, "F(\psi_{I,I})"'] & F(F(I)\otimes I) \arrow[ur, "\varepsilon_{F(I)\otimes I}"'] & \\
& & \\
& F^2(I\otimes I) \arrow[uu, "F(\psi^{-1}_{I,I})"'] \arrow[dr, "\varepsilon_{F(I\otimes I)}"] & \\
F^2(I\otimes I) \arrow[r, "\delta^{-1}_{I\otimes I}"'] \arrow[ur, equal] & F(I\otimes I) \arrow[r, equal] \arrow[u, "\delta_{I\otimes I}"'] & F(I\otimes I) \arrow[uuuu, "\psi^{-1}_{I,I}"']
\end{tikzcd}$$

But the long outside path is composed entirely of isomorphisms. Hence $F(I) \otimes \varepsilon_I$ is invertible, and ε_I is a subunit.

These two constructions are clearly inverse to each other. □

Remark 2.40. Monoidal comonads on $\mathbf{C}$ form a category with morphisms of monoidal comonads [83]. This category is monoidal as a subcategory of $[\mathbf{C}, \mathbf{C}]$. The monoidal unit is the identity comonad $A \mapsto A$. A subunit is a comonad F with a comonad morphism $\lambda\colon F \Rightarrow 1_{\mathbf{C}}$ whose comultiplication is idempotent, and such that $\lambda_A\colon F(A) \to A$

is monic. But by coherence, the latter means that $\varepsilon_I = \lambda_I \colon F(I) \rightarrowtail I$ is monic. It follows that subunits in $\mathbf{C}$ also correspond bijectively to subunits in $[\mathbf{C}, \mathbf{C}]$ in the same sense as Definition 2.1, though we have not strictly defined these since the latter category is not braided. See also [15, Remark 2.3].

It also follows that restrictions monads automatically satisfy the Frobenius law $\delta^{-1}F \circ F\delta = F\delta^{-1} \circ \delta F$ [41], matching the viewpoint in [44].

The second characterisation of subunits s we will give is in terms of the subcategory $\mathbf{C}|_s$.

Definition 2.41. Let $\mathbf{C}$ be a monoidal category. A *monocoreflective tensor ideal* is a full replete subcategory $\mathbf{D}$ such that:

- if $A \in \mathbf{C}$ and $B \in \mathbf{D}$, then $A \otimes B \in \mathbf{D}$;
- the inclusion $F \colon \mathbf{D} \hookrightarrow \mathbf{C}$ has a right adjoint $G \colon \mathbf{C} \to \mathbf{D}$;
- the component of the counit at the tensor unit $\varepsilon_I \colon F(G(I)) \to I$ is monic;
- $F(B) \otimes \varepsilon_I$ is invertible for all $B \in \mathbf{D}$.

Proposition 2.42. *Let* $\mathbf{C}$ *be a firm category. There is a bijection between* $\mathrm{ISub}(\mathbf{C})$ *and the set of monocoreflective tensor ideals of* $\mathbf{C}$.

Proof. A subunit s corresponds to $\mathbf{C}|_s$, and a monocoreflective tensor ideal $\mathbf{D}$ corresponds to ε_I. First notice that $\mathbf{C}|_s$ is indeed a monocoreflective tensor ideal by Proposition 2.30. Starting with $s \in \mathrm{ISub}(\mathbf{C})$ ends up with $s \circ \lambda \colon I \otimes S \rightarrowtail I$, which equals s qua subobject. Starting with a monocoreflective tensor ideal $\mathbf{D}$ ends up with $\{A \in \mathbf{C} \mid A \otimes \varepsilon_I \text{ is invertible}\}$. We need to show that this equals $\mathbf{D}$. One inclusion is obvious. For the other, let $A \in \mathbf{C}$. If $A \otimes \varepsilon_I \colon A \otimes FG(I) \to A \otimes I$ is invertible, then $A \simeq A \otimes F(G(I))$, and so $A \in \mathbf{D}$ because $\mathbf{D}$ is a tensor ideal. □

We leave open the question of what sort of factorisation systems are induced by monocoreflective tensor ideals [18, 23].

2.4 Simplicity

Localisation in algebra generally refers to a process that adds formal inverses to an algebraic structure [54, Chapter 7]. This section discusses how to localise all subunits in a monoidal category at once, by showing that restriction is an example of localisation in this sense.

Definition 2.43. Let $\mathbf{C}$ be a category and Σ a collection of morphisms in $\mathbf{C}$. A *localisation of* $\mathbf{C}$ *at* Σ is a category $\mathbf{C}[\Sigma^{-1}]$ and a functor $Q\colon \mathbf{C} \to \mathbf{C}[\Sigma^{-1}]$ such that:

- $Q(f)$ is an isomorphism for every $f \in \Sigma$;
- for any functor $R\colon \mathbf{C} \to \mathbf{D}$ such that $R(f)$ is an isomorphism for all $f \in \Sigma$, there exists a functor $\overline{R}\colon \mathbf{C}[\Sigma^{-1}] \to \mathbf{D}$ and a natural isomorphism $\overline{R} \circ Q \simeq R$;

$$\begin{tikzcd} \mathbf{C} \arrow[r, "Q"] \arrow[dr, "R"'] & \mathbf{C}[\Sigma^{-1}] \arrow[d, dashed] \arrow[dl, phantom, "\simeq"] \\ & \mathbf{D} \end{tikzcd}$$

- precomposition $(-) \circ Q\colon \big[\mathbf{C}[\Sigma^{-1}], \mathbf{D}\big] \to [\mathbf{C}, \mathbf{D}]$ is full and faithful for every category $\mathbf{D}$.

Proposition 2.44. *Restriction* $\mathbf{C} \to \mathbf{C}|_s$ *at a subunit* s *is a localisation of* $\mathbf{C}$ *at* $\{s \otimes A \mid A \in \mathbf{C}\}$.

Proof. Observe that $S \otimes (-)$ sends elements of Σ to isomorphisms because s is idempotent. Let $R\colon \mathbf{C} \to \mathbf{D}$ be any functor making $R(s \otimes A)$ an isomorphism for all $A \in \mathbf{C}$. Define $\overline{R}\colon \mathbf{C}|_s \to \mathbf{D}$ by $A \mapsto R(A)$ and $f \mapsto R(f)$. Then

$$\eta_A = R(\rho_A) \circ R(s \otimes A)\colon R(s \otimes A) \to R(A)$$

is a natural isomorphism. It is easy to check that precomposition with restriction is full and faithful. □

The above universal property concerns a single subunit. We now move to localising all subunits simultaneously.

Definition 2.45. A monoidal category is *simple* when it has no subunits but I.

In the words of Proposition 2.42, a category is simple when it has no proper monocoreflective tensor ideals. Let us now show how to make a category simple.

Proposition 2.46. *If* $\mathbf{C}$ *is a firm category, then there is a universal simple category* $\mathrm{Loc}(\mathbf{C})$ *with a monoidal functor* $\mathbf{C} \to \mathrm{Loc}(\mathbf{C})$*: any a monoidal functor* $F\colon \mathbf{C} \to \mathbf{D}$ *into a simple category* $\mathbf{D}$ *factors through it via a unique monoidal functor* $\mathrm{Loc}(\mathbf{C}) \to \mathbf{D}$.

$$\begin{tikzcd}
\mathbf{C} \arrow[rr] \arrow[drr, "F"'] & & \mathrm{Loc}(\mathbf{C}) \arrow[d, dashed] \\
& & \mathbf{D}
\end{tikzcd}$$

Proof. We proceed by formally inverting the collection of morphisms

$$\Sigma = \{\lambda_A \circ (s \otimes A) \mid A \in \mathbf{C}, s \in \mathrm{ISub}(\mathbf{C})\} \cup \{A \mid A \in \mathbf{C}\}$$

To show that the localisation $\mathbf{C}[\Sigma^{-1}]$ of Σ exists we will show that Σ admits *a calculus of right fractions* [33]. Firstly, Σ contains all identities and is closed under composition, since the composition of $\lambda_A \circ (A \otimes t)$ and $\lambda_{A\otimes T} \circ (A \otimes T \otimes s)$ is simply $\lambda_A \circ (A \otimes (s \wedge t))$. It remains to show that:

- for morphisms $s\colon A \to C$ in Σ and $f\colon B \to C$ in $\mathbf{C}$, there exist morphisms $t\colon P \to B$ in Σ and $g\colon P \to A$ in $\mathbf{C}$ such that $g \circ s = t \circ f$;

$$\begin{tikzcd}
\bullet \arrow[r, dashed, "g"] \arrow[d, dashed, "\Sigma \ni t"'] & \bullet \arrow[d, "s \in \Sigma"] \\
\bullet \arrow[r, "f"'] & \bullet
\end{tikzcd}$$

- if a morphism $t\colon C \to D$ in Σ and $f, g\colon B \to C$ in $\mathbf{C}$ satisfy $t \circ f = t \circ g$, then $f \circ s = g \circ s$ for some $s\colon A \to B$ in Σ.

It suffices to merely consider $\{\lambda_A \circ (s \otimes A) \mid A \in \mathbf{C}, s \in \mathrm{ISub}(\mathbf{C})\}$ by [31, Remark 3.1]. The first, also called the *right Ore condition*, is satisfied by bifunctoriality of the tensor:

$$\begin{tikzcd}
S \otimes A \arrow[r, dashed, "S \otimes f"] \arrow[d, dashed, "s \otimes A"'] & S \otimes B \arrow[d, "s \otimes B"] \\
I \otimes A \arrow[r, "I \otimes f"'] \arrow[d, dashed, "\rho_A"'] & I \otimes B \arrow[d, "\rho_B"] \\
A \arrow[r, "f"'] & B
\end{tikzcd}$$

For the second, suppose that $(s \otimes B) \circ f = (s \otimes B) \circ g$. Then applying $S \otimes (-)$ and using that $S \otimes s$ is invertible, it follows that $S \otimes f = S \otimes g$. But then

$$\begin{aligned} f \circ \lambda_A \circ (s \otimes A) &= \lambda_{SB} \circ (s \otimes S \otimes B) \circ (S \otimes f) \\ &= \lambda_{SB} \circ (s \otimes S \otimes B) \circ (S \otimes g) = g \circ \lambda_A \circ (s \otimes A), \end{aligned}$$

so the second requirement is satisfied. As a result, $\mathbf{C}[\Sigma^{-1}]$ exists; an easy constuction may be found in [31]. It satisfies the universal property of localisation on the nose. Moreover, the functor $\mathbf{C} \to \mathrm{Loc}(\mathbf{C})$ is monoidal because the class Σ is closed under tensoring with objects of $\mathbf{C}$ by construction [23, Corollary 1.4]. Finally, notice that $\mathrm{Loc}(\mathbf{C})$ is simple by construction. □

2.5 Support

So far we have focused on the spatial structure encoded within the tensor unit. This section investigates how this spatial structure influences arbitrary objects and morphisms in a monoidal category. We will show that Definition 2.26 above gives a well-behaved notion of support, indicating 'where morphisms can happen', that has an appropriate universal property.

When a morphism f restricts to a given subunit s, we might also say that f 'has support in' s. We can use the example of Hilbert modules for the intuition behind this interpretation. Indeed, a subunit $s : C_0(U) \to C_0(X)$ in this setting can be seen as taking functions that are only defined over U and extending them to functions defined over the whole space X by setting them to be zero outside of U. Thus, we can think of the following diagram as saying that, if f can be seen as the result of restricting to region S and then extending by zeroes, then f could only have been non-zero within S to begin with.

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, dashed] & B \\ S \otimes B \arrow[r, "s \otimes B"'] & I \otimes B \arrow[u, "\lambda_B"'] \end{tikzcd}$$

It is natural to assume that each morphism in our category comes with a canonical least subunit to which it restricts, which we may call its support. This is the case in a topos, for example, but in general requires extra structure.

Write $\mathbf{C}^{\preceq}$ for the braided monoidal category whose objects are morphisms $f \in \mathbf{C}$, with $f \otimes g$ defined as in $\mathbf{C}$, tensor unit I, and a unique morphism $f \to g$ whenever $(g \text{ restricts to } s) \implies (f \text{ restricts to } s)$.

Definition 2.47. A *support datum* on a firm category $\mathbf{C}$ is a functor $F\colon \mathbf{C}^{\preceq} \to L$ into a complete lattice L satisfying

$$F(f) = \bigwedge \left\{F(s)\colon s \in \mathrm{ISub}(\mathbf{C}) \mid f \text{ restricts to } s\right\} \tag{2.2}$$

for all morphisms f of $\mathbf{C}$. A *morphism of support data* $F \to F'$ is one of complete lattices $G\colon L \to L'$ with $G \circ F = F'$.

Lemma 2.48. *If $F\colon \mathbf{C}^{\preceq} \to L$ is a support datum, and f, g morphisms in $\mathbf{C}$:*

- *$F(f) = \bigwedge\{F(A) \mid A \in \mathbf{C}, f$ factors through $A\}$;*
- *$F(f \otimes g) \leq F(f) \wedge F(g)$ for all f, g; so F is colax monoidal.*

This notion of support via objects is similar to that of [6, 58, 50].

Proof. For the first statement, it suffices to show that f restricts to a subunit s iff it factors through some object A which does. But if f factors through A then $f = g \circ A \circ h$ for some g, h and so if A restricts to s so does f. Conversely if $f\colon B \to C$ restricts to s it factors over $S \otimes C$, which always restricts to s.

For the second statement, Note that $F(I) \leq 1$ always, so colax monoidality reduces to the rule above. But if f restricts to s then so does $f \otimes g$. Hence $F(f \otimes g) \leq F(f)$, and $F(f \otimes g) \leq F(g)$ similarly. □

Most features of support data follow from the associated map $\mathrm{ISub}(\mathbf{C}) \to L$.

Proposition 2.49. *Let $\mathbf{C}$ be a firm category and L a complete lattice. Specifying a support datum $F\colon \mathbf{C}^{\preceq} \to L$ is equivalent to specifying a monotone map $\mathrm{ISub}(\mathbf{C}) \to L$.*

Proof. In $\mathbf{C}^{\preceq}$ there is a morphism $s \to t$ between subunits s and t precisely when $s \leq t$. Hence any support datum restricts to a monotone map $\mathrm{ISub}(\mathbf{C}) \to L$.

Conversely, let F be such a map and extend it to arbitrary morphisms by (2.2). Both definitions of F agree on subunits s since a subunit restricts to another one t precisely when $s \leq t$, so that $F(s) = \bigwedge\{F(t) \mid s \leq t\}$. Finally, for functoriality suppose there exists a morphism $f \to g$ in $\mathbf{C}^{\preceq}$. If this holds then whenever g restricts to s then so does f, so that $F(f) \leq F(g)$. □

This observation provides examples of support data. Recall that the free complete lattice on a semilattice L is given by its collection $D(L)$ of downsets $U = \downarrow U \subseteq L$ under inclusion, via the embedding $x \mapsto \downarrow x$ [48, II.1.2].

Proposition 2.50. *Any firm category* $\mathbf{C}$ *has a canonical support datum, valued in* $D(\mathrm{ISub}(\mathbf{C}))$*, given by*

$$\mathrm{supp}_0(f) = \{s \in \mathrm{ISub}(\mathbf{C}) \mid f \text{ restricts to } t \implies s \leq t\}. \tag{2.3}$$

Moreover, supp_0 *is initial: any support datum factors through it uniquely.*

$$\begin{tikzcd} \mathbf{C}^{\leq} \arrow[r, "\mathrm{supp}_0"] \arrow[dr, "F"'] & D(\mathrm{ISub}(\mathbf{C})) \arrow[d, dashed] & \{s_i\} \arrow[d, mapsto] \\ & L & \bigvee F(s_i) \end{tikzcd}$$

This generalises [6, 7, 8] from triangulated categories to firm ones.

Proof. Extend the embedding $L \to D(L)$ to a support datum via Proposition 2.49. Initiality is immediate by freeness of $D(L)$, with (2.3) coming from the description of meets in terms of joins in a complete lattice. □

Rather than require extra data, it would be desirable to define support internally to the category. If $\mathbf{C}$ has the property that $\mathrm{ISub}(\mathbf{C})$ is already a complete lattice (or frame), then it indeed comes with a support datum given by the identity on $\mathrm{ISub}(\mathbf{C})$. We may then define *the support* of a morphism as

$$\mathrm{supp}(f) = \bigwedge \left\{ s \in \mathrm{ISub}(\mathbf{C}) \mid f \text{ restricts to } s \right\}.$$

Note that $\mathrm{supp}(f) = \bigvee \mathrm{supp}_0(f)$. It therefore follows from Proposition 2.50 that supp also has a universal property: if $\mathrm{ISub}(\mathbf{C})$ is already a complete lattice, any support datum F factors through supp via a semilattice morphism. Therefore, in the case of a topos, $\mathrm{supp}(A)$ is the factorisation of a morphism $A \to 1$ into a strong epimorphism and a monomorphism.

Example 2.51. Let L be a frame and consider $\mathrm{Sh}(L)$. A morphism $f \colon A \Rightarrow B$ has $\mathrm{supp}_0(f) = \downarrow\{t \mid A(t) \neq \emptyset\}$, and $\mathrm{supp}(f) = \bigwedge\{s \mid A(s) \neq \emptyset\}$.

Example 2.52. In $\mathbf{Hilb}_{C_0(X)}$ the collection of subunits forms a frame, and each morphism $f \colon A \to B$ has $\mathrm{supp}(f) = C_0(U_f)$, where

$$U_f = \{x \in X \mid \langle f(a) \mid f(a) \rangle(x) \neq 0 \text{ for some } a \in A\}.$$

Letting L be the totally ordered set of cardinals below $|X|$, we may define another support datum by $F(f) = |U_f| \in L$.

In the remaining sections, we turn to categories coming with such an intrinsic spatial structure. First, the following example shows that, even in case $\mathrm{ISub}(\mathbf{C})$ is a frame, our notion of support differs from that of [6, Definition 3.1(SD5)] and [58, Definition 3.2.1(5)]: without further assumptions, a support datum is only colax monoidal.

Example 2.53. There is a firm category $\mathbf{C}$ for which $\mathrm{ISub}(\mathbf{C})$ is a frame but

$$\mathrm{supp}(f) \otimes \mathrm{supp}(g) \neq \mathrm{supp}(f \otimes g).$$

Proof. Let Q be the commutative unital quantale with elements $0 \leq \varepsilon \leq 1$, with unit 1 and satisfying $0 = 0 \cdot 0 = 0 \cdot \varepsilon = \varepsilon \cdot \varepsilon$. Then the frame of subunits is $\mathrm{ISub}(Q) = \{0, 1\}$, and ε satisfies $\mathrm{supp}(\varepsilon) = 1$ whereas $\mathrm{supp}(\varepsilon \cdot \varepsilon) = 0$. □

2.6 Locale-based categories

In our main examples, the subunits satisfy extra properties over being a mere semilattice, and they interact universally with the rest of the category. First, they often satisfy the following property.

Definition 2.54. A category is *stiff* when it is braided monoidal and

$$\begin{tikzcd} S \otimes T \otimes X \arrow[r, tail, "s \otimes T \otimes X"] \arrow[d, tail, "S \otimes t \otimes X"'] \arrow[dr, phantom, very near start, "\lrcorner"] & T \otimes X \arrow[d, tail, "t \otimes X"] \\ S \otimes X \arrow[r, tail, "s \otimes X"'] & X \end{tikzcd} \tag{2.4}$$

is a pullback of monomorphisms for all objects X and subunits s, t.

Any stiff category is firm: take $X = I$ and recall that pullbacks of monomorphisms are monomorphisms. More strongly, subunits often come with joins satisfying the following.

Definition 2.55. Let $\mathbf{C}$ be a braided monoidal category. We say that $\mathbf{C}$ has *universal finite joins* of subunits when it has an initial object 0 whose morphism $0 \longrightarrow I$ is monic, with $X \otimes 0 \simeq 0$ for all objects X, and $\mathrm{ISub}(\mathbf{C})$ has finite joins such that each diagram

$$\begin{tikzcd} S \otimes T \otimes X \arrow[r, tail] \arrow[d, tail] \arrow[dr, phantom, very near start, "\lrcorner"] \arrow[dr, phantom, very near end, "\ulcorner"] & T \otimes X \arrow[d, tail] \\ S \otimes X \arrow[r, tail] & (S \vee T) \otimes X \end{tikzcd} \tag{2.5}$$

is both a pullback and pushout of monomorphisms, where each morphism is the obvious inclusion tensored with X as in (2.4).

Lemma 2.56. *Let* $\mathbf{C}$ *be braided monoidal with universal finite joins of subunits. Then* $\mathbf{C}$ *is stiff and* $\mathrm{ISub}(\mathbf{C})$ *is a distributive lattice with least element* 0.

Proof. For stiffness, take $t = 1_I$ to see that each morphism $s \otimes X$ is monic. Then since $(s \vee t) \otimes X$ is monic it follows easily that each diagram (2.4) is a pullback. By assumption $0 \to I$ is indeed a subunit. Finally it follows from (2.5) with $X = R$ that subunits R, S, T satisfy $(S \vee T) \wedge R = (S \wedge R) \vee (T \wedge R)$. □

Example 2.57. Any coherent category $\mathbf{C}$ forms a cartesian monoidal category with universal finite joins of subunits.

Proof. Each partial order $\mathrm{Sub}(A)$ is a distributive lattice, and for subobjects $S, T \rightarrowtail A$ each diagram (2.5) with $\wedge$ replacing $\otimes$ and $X = 1$ is indeed both a pushout and pullback [49, A1.4.2, A1.4.3]. Moreover in such a category each functor $X \times (-)$ preserves these pullbacks, since limits commute with limits, and preserves finite joins and hence these pushouts since each functor $(\pi_2)^* \colon \mathrm{Sub}(A) \to \mathrm{Sub}(X \times A)$ does so by coherence of $\mathbf{C}$. □

To obtain arbitrary joins of subunits from finite ones, it will suffice to also have the following. Recall that a subset U of a partially ordered set is (upward) *directed* when any $a, b \in U$ allow $c \in U$ with $a \leq c \geq b$. A *preframe* is a semilattice in which every directed subset has a supremum, and finite meets distribute over directed suprema.

By a *directed colimit of subunits* we mean a colimit of a diagram $D \colon \mathbf{J} \to \mathbf{C}$, for which $\mathbf{J}$ is a directed poset, all of whose arrows are inclusions $S_i \rightarrowtail S_j$ between a collection of subunits $s_i \colon S_i \to I$. In particular D has a cocone given by these subunits, inducing a morphism $\mathrm{colim}\, D \to I$ if a colimit exists.

Definition 2.58. A stiff category $\mathbf{C}$ has *universal directed joins* of subunits when it has directed colimits of subunits, each of whose induced arrow $\mathrm{colim}\, S \to I$ is again a subunit, and these colimits are preserved by each functor $X \otimes (-)$.

Lemma 2.59. *If a stiff category* $\mathbf{C}$ *has universal directed joins of subunits, then* $\mathrm{ISub}(\mathbf{C})$ *is a preframe.*

Proof. Any directed subset $U \subseteq \mathrm{ISub}(\mathbf{C})$ induces a diagram $U \to \mathbf{C}$, and its colimit is by assumption a subunit which is easily seen to form a supremum of U. Taking X to be a subunit shows that $\wedge$ distributes over directed suprema. □

Example 2.60. Any preframe L, regarded as a monoidal category under $(\wedge, 1)$, has universal directed joins.

The rest of this section shows that the subunits of a category have a spatial nature when it has both types of universal joins above. We unify Definitions 2.55 and 2.58 as follows. Let $\mathbf{C}$ be a braided monoidal category and $U \subseteq \mathrm{ISub}(\mathbf{C})$ a family of subunits. For any object X, write $D(U,X)$ for the diagram of objects $S \otimes X$ for $s \in U$ and all morphisms $f : S \otimes X \to T \otimes X$ satisfying $(t \otimes X) \circ f = s \otimes X$. If $\mathbf{C}$ is stiff, there is a unique such f for s and t.

$$\begin{tikzcd} S \otimes X \arrow[rr, dashed] \arrow[dr, "s \otimes X"'] & & T \otimes X \arrow[dl, "t \otimes X"] \\ & X & \end{tikzcd}$$

Call such a set U of subunits *idempotent* when $U = U \otimes U := \{s \wedge t \mid s, t \in U\}$.

Definition 2.61. A category $\mathbf{C}$ is *locale-based* when it is stiff, $\mathrm{ISub}(\mathbf{C})$ is a frame, and the canonical maps $S \otimes X \to (\bigvee U) \otimes X$ form a colimit of $D(U,X)$ for each idempotent $U \subseteq \mathrm{ISub}(\mathbf{C})$ and $X \in \mathbf{C}$.

Let us now see how this combines our earlier notions. In any poset P, an *ideal* is a downward closed, upward directed subset. Let us call a subset $U \subseteq P$ *finitely bounded* when it has a finite non-empty subset $T = \{x_1, \ldots, x_n\}$ of elements such that $u \leq x$ for every $x \in T$ and every $u \in U - T$. If U is downward closed then equivalently it is finitely generated: $U = \downarrow\{x_1, \ldots, x_n\}$.

Proposition 2.62. *A category* $\mathbf{C}$ *has universal finite (directed) joins if and only if* $\mathrm{ISub}(\mathbf{C})$ *has finite (directed) joins, and* $D(U,X)$ *has colimit* $S \otimes X \to (\bigvee U) \otimes X$ *for each idempotent* $U \subseteq \mathrm{ISub}(\mathbf{C})$ *that is finitely bounded (directed).*

Proof. First consider finite joins. A colimit of $D(\emptyset, X)$ is precisely an initial object and the conditions on 0 in both cases are equivalent to $0 \to I$ being a subunit with $0 \otimes X \simeq 0$ for all X. Moreover in any stiff category it is easy to see that cocones over the top left corner of (2.5) correspond to those over $D(\downarrow\{s,t\}, X)$. (See also Lemma 2.67 below.) Hence the properties above provide each diagram with a colimit $(S \vee T) \otimes X$, and so $\mathbf{C}$ with universal finite joins.

Conversely, suppose that $\mathbf{C}$ has universal finite joins. For any idempotent U we claim that any cocone c_s over $D(U,X)$ extends to one over $D(V,X)$, where

$V = \{s_1 \vee \cdots \vee s_n \mid s_i \in U\}$. Indeed for any $s, t \in U$ the following diagram commutes, giving $c_{s \vee t}$ as the unique mediating morphism.

$$
\begin{tikzcd}
S \otimes T \otimes X \arrow[r, tail] \arrow[d, tail] & T \otimes X \arrow[d] \arrow[ddr, bend left, "c_t"] & \\
S \otimes X \arrow[r] \arrow[drr, bend right, "c_s"'] & (S \vee T) \otimes X \arrow[ul, phantom, "\ulcorner", very near start] \arrow[dr, dashed, "c_{s \vee t}"] & \\
 & & C
\end{tikzcd}
$$

Similarly define morphisms $c_{s_1 \vee \cdots \vee s_n}$ for arbitrary elements of V; these form a cocone. Hence $\operatorname{colim} D(U, X) = \operatorname{colim} D(V, X)$. But if U is bounded by some $s_1, \ldots, s_n$ then clearly $\operatorname{colim} D(V, X) = (s_1 \vee \cdots \vee s_n) \otimes X$ and we are done.

Next, consider directed joins. Let D be a directed diagram of inclusions between elements of $U \subseteq \mathrm{ISub}(\mathbf{C})$. Then U must be directed and therefore $V = \{s_1 \wedge \cdots \wedge s_n \mid s_i \in U\}$ is idempotent and directed. Moreover, for each object X, any cocone c_s over $D \otimes X$ extends to one over $D(V, X)$: for any $s \in V$, let $s \leq t \in U$ and set $c_s = c_t \circ (x \otimes 1_X)$ where $x \colon S \to T$ is the inclusion. Since $R = \bigvee V$ has $R \otimes X = \operatorname{colim} D(V, X)$ then $R \otimes X = \operatorname{colim}(D \otimes X)$ as required.

Conversely, suppose $\mathbf{C}$ has universal directed joins. Then $\mathrm{ISub}(\mathbf{C})$ is a preframe by Lemma 2.59. If $U \subseteq \mathrm{ISub}(\mathbf{C})$ is directed and idempotent then for each X we have $R \otimes X = \operatorname{colim}\big(D(U, I) \otimes X\big)$, where $R = \bigvee U$. But any cocone over $D(U, X)$ certainly also forms one over $D(U, I) \otimes X$, and so $R \otimes X = \operatorname{colim} D(U, X)$ also. □

Corollary 2.63. *A category is locale-based if and only if it has universal finite and directed joins of subunits.*

Proof. Proposition 2.62 proves one direction. In the other direction, suppose $\mathbf{C}$ has universal finite and directed joins of subunits. Then $\mathrm{ISub}(\mathbf{C})$ is a frame by Lemmas 2.56 and 2.59, since a poset is a frame precisely when it is a preframe and a distributive lattice. Let $U \subseteq \mathrm{ISub}(\mathbf{C})$ be idempotent. Then $V = \{s_1 \vee \cdots \vee s_n \mid s_i \in U\}$ is idempotent by distributivity, as well as directed, so that $\operatorname{colim} D(V, X) = (\bigvee V) \otimes X$ exists for any X. But $\operatorname{colim} D(U, X) = \operatorname{colim} D(V, X)$ as in the proof of Proposition 2.62. □

The previous corollary justifies saying that a category simply *has universal joins* of subunits when it is locale-based. The rest of this section shows that our main examples are locale-based.

Example 2.64. Any commutative unital quantale Q is locale-based when regarded as a category as in Example 2.14; in particular so is any frame under tensor $\wedge$. Indeed that example showed that $\mathrm{ISub}(Q)$ is a frame, and for any $U \subseteq \mathrm{ISub}(Q)$ and $x \in Q$ we have $\operatorname{colim} D(U,x) = \bigvee_{s\in U} sx = (\bigvee_{s\in U} s)x$.

Example 2.65. Any cocomplete Heyting category $\mathbf{C}$ is locale-based under cartesian products. This includes all cocomplete toposes, such as Grothendieck toposes.

Proof. Since a Heyting category is coherent, it has universal finite joins by Example 2.57, with each change of base functor having a right adjoint and so preserving arbitrary joins of subobjects. In any cocomplete regular category with this property, for any directed diagram D and any cocone C over D all of whose legs are monic, the induced map $\operatorname{colim} D \to C$ is again monic [36, Corollary II.2.4]. Hence whenever U is directed, so is each map $\operatorname{colim} D(U,X) \to X$, ensuring that $\operatorname{colim} D(U,X) = \bigvee_{s\in U} s \times X$ is in $\mathrm{Sub}(X)$. Since each functor $X \times (-)$ now preserves arbitrary joins of subobjects furthermore $\bigvee_{s\in U} s \times X = \operatorname{colim} D(U,I) \times X$, establishing universal directed joins. □

Next we consider Hilbert modules. In general $\mathbf{Hilb}_{C_0(X)}$ is finitely cocomplete but not cocomplete, and so lacks directed colimits by [66, IX.1.1]; this follows from [3, Example 2.3 (9)] by taking X to be trivial and so reducing to the category of Hilbert spaces and contractive linear maps. Nonetheless, we have the following.

Example 2.66. $\mathbf{Hilb}_{C_0(X)}$ is locale-based.

Proof. Throughout this proof we again identify $C_0(U)$ with the submodule (2.1) of $C_0(X)$, and identify the module $A \otimes C_0(U)$ with $A|_U$, for open $U \subseteq X$.

First let us show that $\mathbf{Hilb}_{C_0(X)}$ has universal finite joins of subunits. For open subsets $U, V \subseteq X$, and any Hilbert $C_0(X)$-module A, consider the diagram of inclusions between $A|_{U\cap V}$, $A|_U$, $A|_V$ and $A|_{U\cup V}$. It is easily seen to be a pullback, since $A|_{U\cap V} = A|_U \cap A|_V$ as subsets of A. We verify that it is also a pushout. Since any morphism $A_{U\cup V} \to B$ restricts to $C_0(U\cup V)$, it suffices to assume that $X = U \cup V$. We claim that

$$C_0(U) + C_0(V) = \{g_U + g_v \in C_0(X) \mid g_U \in C_0(U), g_V \in C_0(V)\}$$

is a dense submodule of $C_0(X)$. To see this, let $g \in C_0(X)$ and $\varepsilon > 0$, and K be compact with $|g(x)| \geq \epsilon \implies x \in K$. Urysohn's lemma for locally compact Hausdoff spaces [79, 2.12] produces $h \in C_0(U)$ such that $|h(x)| \leq |g(x)|$ for $x \in U$ and $h(x) = g(x)$ for $x \in K \setminus V$. Then $|(g-h)(x)| \geq 2\varepsilon \implies x \in L$ for some compact $L \subseteq K \cap V$. Again there is $k \in C_0(V)$ with $|k(x)| \leq |g(x)|$ for all $x \in V$ and $k(x) = (g-h)(x)$ for $x \in L$. By construction $\|g-h-k\| \leq 4\varepsilon$, establishing the claim. It follows also that

$$A|_U + A|_V = \{a_U + a_V \mid a_U \in A|_U, a_V \in A|_V\}$$

is dense in A, since $A \cdot C_0(X) = \{a \cdot g \mid g \in C_0(X)\}$ is so too [62, p5].

Now suppose $f_U \colon A|_U \to B$ and $f_V \colon A|_V \to B$ agree on $A|_{U\cap V}$. Then for $a = a_U + a_V$ with $a_U \in A|_U$ and $a_V \in A|_V$, the assignment

$$f(a) = f_U(a_U) + f_V(a_V)$$

is a well-defined A-linear map. Hence it extends to a unique map $f \colon A \to B$ which is by definition the unique factorisation of f_U and f_V through the diagram.

Now we must check that f is contractive when f_U and f_V are. Let $x \in X$, and without loss of generality say $x \in U$. Urysohn's lemma again produces $g \in C_0(U)$ with $g(x) = 1 = \|g\|$. Now $a \cdot g \in A|_U$ for any $a \in A$. So, writing $|a|^2(x)$ for $|\langle a \mid a\rangle(x)|$, we find

$$|f(a)|(x) = |f(a) \cdot g|(x) \leq \|f(a) \cdot g\| = \|f_U(a) \cdot g\| \leq \|a\|\|g\| \leq \|a\|$$

using $\|f_U\| \leq 1$. Since x was arbitrary, also $\|f\| \leq 1$.

Next, let us consider universal directed joins of subunits. For this, let W be a directed family of open sets in X; again it suffices to assume $X = \bigcup W$. We claim that

$$\bigcup_{U\in W} C_0(U) = \{g \in C_0(X) \mid g \in C_0(U) \text{ for some } U \in W\}$$

is a dense submodule of $C_0(X)$. Again let $g \in C_0(X)$ and $\varepsilon > 0$, and let K be compact with $|g(x)| \geq \epsilon \implies x \in K$. Since K is compact and W is directed, $K \subseteq U$ for some $U \in W$. Urysohn again provides $h \in C_0(U)$ with $|h(x)| \leq |g(x)|$ for all $x \in U$ and $h(x) = g(x)$ for $x \in K$. Then $|g - h|(x) \leq |g(x)| + |h(x)| \leq 2\varepsilon$ for $x \in X \setminus K$ and so, since g and h agree on K, we have $\|g - h\| \leq 2\varepsilon$, establishing the claim. Similarly, for any Hilbert module A, since $A \cdot C_0(X)$ is dense in A, so is $\bigcup_{U\in W} A|_U$.

Finally, let $f_U \colon A|_U \to B$ be a cocone over $D(W, A)$. It suffices to show that there is a unique $f \colon A \to B$ with $f(a) = f_U(a)$ for all $a \in A|_U$. But any $a \in A$ has $a = \lim(a_n)_{n=1}^{\infty}$ with each $a_n \in A|_{U_n}$ for some U_n. By directedness we may assume $U_n \subseteq U_{n+1}$ for all n. Then $f \colon A \to B$ must satisfy $f(a) = \lim f_{U_n}(a_n)$, making f unique. Additionally, this limit is always well-defined since a_n is a Cauchy sequence and so for $n \leq m$:

$$\|f_{U_n}(a_n) - f_{U_m}(a_m)\| = \|f_{U_m}(a_n - a_m)\| \leq \|a_n - a_m\|$$

and $f_{U_n}(a_n)$ is also a Cauchy sequence. Clearly f is A-linear and $\|f\| \leq 1$. □

2.7 Universal joins from colimits

This section characterises each of the notions of universal joins purely categorically, without order-theoretic assumptions on $\mathrm{ISub}(\mathbf{C})$. Instead, they will be cast solely in terms of the diagrams $D(U,X)$. When we turn to completions in the next sections, we can therefore use the diagrams $D(U,X)$ themselves as formal joins to add.

Lemma 2.67. *Let* $\mathbf{C}$ *be a stiff category. If* $U \subseteq \mathrm{ISub}(\mathbf{C})$ *is idempotent, then any cocone over* $D(U,X)$ *extends uniquely to one over* $D(\downarrow U, X)$*.*

Therefore, $\mathbf{C}$ has colimits of $D(U,X)$ for all downward-closed $U \subseteq \mathrm{ISub}(\mathbf{C})$ if and only if it has them for idempotent U.

Proof. Let U be idempotent and consider a cocone $c_s \colon S \otimes X \to X$ over $D(U,X)$. Let $r \in \downarrow U$, say $r = s \circ f$ for $s \in U$ and $f \colon R \to S$. Define $c_r = c_s \circ (f \otimes X) \colon R \otimes X \to X$. This is clearly the only possible extension of c_s to $D(\downarrow U, X)$. We will prove that it is a well-defined cocone. Suppose $r' \in \mathrm{ISub}(\mathbf{C})$ satisfies $r' \leq s'$ for $s' \in U$, and $r \otimes X = (r' \otimes X) \circ g$. Then the marked morphism in the following diagram is an isomorphism:

$$\begin{tikzcd}
R \otimes X \arrow[rr, tail, "g"] \arrow[dd, tail] & & R' \otimes X \arrow[dd, tail] \\
& R \otimes R' \otimes X \arrow[ul, tail, "\simeq"', "R \otimes r' \otimes X"] \arrow[ur, tail, "r \otimes R' \otimes X"'] \arrow[d, tail] & \\
S \otimes X \arrow[dr, "c_s"'] & S \otimes S' \otimes X \arrow[l, tail, "S \otimes s' \otimes X"'] \arrow[r, tail, "s \otimes S' \otimes X"] & S' \otimes X \arrow[dl, "c_{s'}"] \\
& X &
\end{tikzcd}$$

The upper triangle and central squares commute trivially. The lower quadrilateral commutes and equals $c_{s \otimes s'}$ because $s \otimes s' \in U$ and c is a cocone. Hence the outer diagram commutes, showing $c_r = c_{r'} \circ g$ as required. In particular, taking $R' = R$ shows that c_r is independent of the choice of s. □

Lemma 2.68. *Let* $\mathbf{C}$ *and* $\mathbf{D}$ *be stiff categories,* $U \subseteq \mathrm{ISub}(\mathbf{C})$ *be idempotent, and* $c_s \colon S \otimes X \to C$ *be a cocone over* $D(U,X)$*. If a functor* $F \colon \mathbf{C} \to \mathbf{D}$ *preserves monomorphisms of the form* $s \otimes X \rightarrowtail X$*, for subunits* s*, and the pullbacks* (2.4)*, then* $F(c_s)$ *is a cocone over* $D\big(F(U), F(X)\big)$*, where* $F(U) = \{F(s) \mid s \in U\}$*.*

Proof. Clearly, if $s \otimes X \leq t \otimes X$ then $F(s \otimes X) \leq F(t \otimes X)$, and $F(c_s)$ respects the inclusion. Conversely, suppose that $F(s \otimes X) \leq F(t \otimes X)$ via some morphism f, and consider the following diagram.

$$\begin{tikzcd}
F(S\otimes T\otimes X) \arrow[r, tail, "F(s\otimes T\otimes X)"] \arrow[d, tail, "F(S\otimes t\otimes X)"'] & F(T\otimes X) \arrow[d, "F(c_t)"] \arrow[ddr, bend left, "F(t\otimes X)"] & \\
F(S\otimes X) \arrow[ur, "f"] \arrow[r, "F(c_s)"'] \arrow[drr, bend right, "F(s\otimes X)"'] & F(C) & \\
& & F(X)
\end{tikzcd}$$

The outer rectangle commutes by bifunctoriality, and $F(t\otimes X)\circ f = F(s\otimes X)$ by assumption. Hence the upper left triangle commutes because $F(t\otimes X)$ is monic by stiffness and the assumption on F. The inner square commutes and is equal to $F(c_{s\otimes t})$ by definition of $D(U,X)$. Since the outer rectangle is a pullback, the leftmost vertical morphism is invertible and hence $F(c_t)\circ f = F(c_s)$. $\square$

Now suppose a diagram $D(U,X)$ has a colimit $c_s^X\colon S\otimes X\to \operatorname{colim} D(U,X)$ for each idempotent $U\subseteq \mathrm{ISub}(\mathbf{C})$ and object X. Then there are two canonical morphisms. First, a mediating map $\operatorname{colim} D(U,I)\to I$ to the cocone $s\colon S\to I$.

$$\begin{tikzcd}
S \arrow[r, "c_s^I"] \arrow[dr, "s"'] & \operatorname{colim} D(U,I) \arrow[d, dashed] \\
& I
\end{tikzcd} \tag{2.6}$$

Second, in a stiff category it follows from applying Lemma 2.68 to $(-)\otimes X$ that there is a unique map making the following triangle commute for all $s\in U$:

$$\begin{tikzcd}
S\otimes X \arrow[r, "c_s^X"] \arrow[dr, "c_s^I\otimes X"'] & \operatorname{colim} D(U,X) \arrow[d, dashed] \\
& (\operatorname{colim} D(U,I))\otimes X
\end{tikzcd} \tag{2.7}$$

If $\mathbf{C}$ has universal joins of U then $\bigvee U = \operatorname{colim} D(U,I)$ and (2.6) is monic, and (2.7) is invertible by definition. We now set out to prove the converse.

Lemma 2.69. *Let $\mathbf{C}$ be a stiff category, and let $U\subseteq \mathrm{ISub}(\mathbf{C})$ be idempotent. Suppose that $D(U,X)$ has a colimit for each object X and that each morphism* (2.7) *is an isomorphism. If the morphism* $\operatorname{colim} D(U,I)\to I$ *of* (2.6) *is monic, then it is a subunit.*

Proof. Write s_U for this morphism, which is monic by assumption. For each $s \in U$, we claim $S \otimes s_U \colon S \otimes \operatorname{colim} D(U,I) \to S$ is an isomorphism. It is monic because

$$s_U \circ c_s \circ (S \otimes s_U) = s \otimes s_U = s_U \circ \big(s \otimes \operatorname{colim} D(U,I)\big)$$

where s_U and $s \otimes \operatorname{colim} D(U,I)$ are monic by stiffness. But it is also split epic since $(S \otimes s_U) \circ (S \otimes c_s) = S \otimes s$ is an isomorphism.

Now since $s \circ (S \otimes s_U) = s_U \circ (s \otimes \operatorname{colim} D(U,I))$, bifunctoriality of $\otimes$ shows that for all $s,t \in U$:

$$s \otimes \operatorname{colim} D(U,I) \leq t \otimes \operatorname{colim} D(U,I) \qquad \Longleftrightarrow \qquad s \leq t$$

This gives an isomorphism of diagrams

$$S \otimes s_U \colon S \otimes \operatorname{colim} D(U,I) \to S$$

from $D\big(U, \operatorname{colim} D(U,I)\big)$ to $D(U,I)$. Writing $c_s \colon S \to \operatorname{colim} D(U,I)$ for the latter colimit, $c_s \otimes \operatorname{colim} D(U,I)$ is a colimit for the former by assumption. Hence the unique map making the following square commute

$$\begin{tikzcd}
S \otimes \operatorname{colim} D(U,I) \arrow[r, "S \otimes s_U"] \arrow[d, "c_s \otimes \operatorname{colim} D(U,I)"'] & S \arrow[d, "c_s"] \\
\operatorname{colim} D(U,I) \otimes \operatorname{colim} D(U,I) & \operatorname{colim} D(U,I) \arrow[l]
\end{tikzcd}$$

is invertible. But this map is just $\operatorname{colim} D(U,I) \otimes s_U$, so s_U is a subunit. $\square$

We can now characterise locale-based categories purely categorically.

Theorem 2.70. *A stiff category* $\mathbf{C}$ *has universal (finite, directed) joins if and only if for each idempotent (and finitely bounded, directed)* $U \subseteq \mathrm{ISub}(\mathbf{C})$*:*

- *the diagram* $D(U,X)$ *has a colimit;*
- *the canonical morphism* (2.6) *is monic;*
- *the canonical morphism* (2.7) *is invertible.*

Proof. The conditions are clearly necessary, as already discussed. Conversely, suppose that they hold and let $U \subseteq \mathrm{ISub}(\mathbf{C})$ be as above. Lemma 2.67 lets us assume $U = {\downarrow} U$. Then $s_U \colon \operatorname{colim} D(U,I) \to I$ is a subunit by Lemma 2.69, and by definition $s \leq s_U$ for all $s \in U$. Now suppose that t is also an upper bound in $\mathrm{ISub}(\mathbf{C})$ of all $s \in U$. Then the inclusions $i_{s,t}\colon S \to T$ form a cocone over $D(U,I)$. Hence there is a unique mediating map $f\colon \operatorname{colim} D(U,I) \to T$ with $i_{s,t} = f \circ c_s^I$ for all $s \in U$. But then

$$t \circ f \circ c_s^I = t \circ i_{s,t} = s = s_U \circ c_s^I$$

for all $s \in U$. Because the c_s^I are jointly epic, $t \circ f = s_U$, so that $s_U \leq t$. Therefore indeed $\operatorname{colim} D(U,I) = \bigvee U$. Thus universal finite or directed joins follow by Proposition 2.62, and so arbitrary ones by Corollary 2.63. □

2.8 Completions

Our goal for this section is to embed a stiff category $\mathbf{C}$ into one with any given kind of universal joins of subunits, including a locale-based category. One might think to work with the free cocompletion of $\mathbf{C}$, the category of presheaves $\widehat{\mathbf{C}} = [\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$. Here, $\widehat{\mathbf{C}}$ is endowed with the Day convolution $\widehat{\otimes}$ as tensor; for details see Appendix A. Although $\widehat{\mathbf{C}}$ has a complete lattice of subunits, we will see that it has two problems: it is in general not firm, and it has too many subunits to be the locale-based completion. We will remedy both problems by passing to a full subcategory of so-called broad presheaves.

First, note that any subunit s in a firm category $\mathbf{C}$ induces a subunit

$$s \circ (-)\colon \mathbf{C}(-,S) \to \mathbf{C}(-,I)$$

in $\widehat{\mathbf{C}}$ since the Yoneda embedding is monoidal, full, and faithful, and preserves all limits and hence monomorphisms.

Proposition 2.71. *If $\mathbf{C}$ is a cocomplete regular category, and for all objects A the functors $A \otimes (-)$ preserve colimits, then* $\mathrm{ISub}(\mathbf{C})$ *is a complete lattice. Thus, if $\mathbf{C}$ is any braided monoidal category, then* $\mathrm{ISub}(\widehat{\mathbf{C}})$ *is a complete lattice.*

Proof. In cocomplete regular categories, the subobjects of a fixed object form a complete lattice [12, Proposition 4.2.6]. Explicitly, let $s_i\colon S_i \rightarrowtail I$ be a family of subunits. Choose a coproduct $c_i\colon S_i \to C$. The unique mediating map $C \to I$ factors through a monomorphism $\bigvee s_i\colon S \rightarrowtail I$, which is the supremum.

$$\begin{tikzcd}
S_i \arrow[r, "c_i"] \arrow[ddr, tail, bend right, "s_i"'] & C \arrow[d, dashed, two heads, "e"] & S_j \arrow[l, "c_j"'] \arrow[ddl, tail, bend left, "s_j"] \\
 & S \arrow[d, dashed, tail, "s"] & \\
 & I &
\end{tikzcd}$$

Next we show that $\bigvee s_i$ is a subunit. Let $c = s \circ e \colon C \to I$. We claim that

$$\begin{tikzcd}
C \otimes C \arrow[rr, "C \otimes c"] \arrow[dr, "\simeq"'] & & C \\
& \coprod_i S_i \otimes C \arrow[ur, "\coprod_i (S_i \otimes c)"'] &
\end{tikzcd}$$

is a regular epimorphism. Since colimits commute with colimits, it suffices to check that each $S_i \otimes c$ is a regular epimorphism. But this is so: if $S_i \otimes c = m \circ f$ for some regular epimorphism f and monomorphism m, then $m \circ f \circ (S_i \otimes c_i) = (S_i \otimes c) \circ (S_i \otimes c_i) = S_i \otimes s_i$ is an isomorphism by idempotence of s_i, so that m is split epic as well as monic and hence an isomorphism.

Now the topmost two rectangles in the following diagram commute.

$$\begin{tikzcd}
S_i \arrow[dr, "c_i"'] \arrow[dddr, bend right, "s_i"'] & & & S_i \otimes S_i \arrow[lll, "S_i \otimes s_i"'] \arrow[dl, "c_i \otimes c_i"'] \arrow[dddl, bend left, "s_i \otimes s_i"] \\
& C \arrow[d, "e"'] & C \otimes C \arrow[l, "C \otimes c"'] \arrow[d, "e \otimes e"] & \\
& S \arrow[d, "s"'] & S \otimes S \arrow[l, "\lambda_S \circ (S \otimes s)"'] \arrow[d, "s \otimes s"] & \\
& I & I \otimes I \arrow[l, "\lambda_I"] &
\end{tikzcd}$$

The left and right triangles commute by construction, and the bottom rectangle commutes by bifunctoriality of the tensor and naturality of λ. Because e is a coequaliser, so are $C \otimes e$ and $e \otimes S$, and hence so is $e \otimes e$. Therefore both vertical morphisms factor as regular epimorphisms followed by monomorphisms, and the mediating morphism, which must be $\lambda_S \circ (S \otimes s)$ by uniqueness, is an isomorphism. Thus $S \otimes s$ is an isomorphism, as required.

The second statement now follows, because $\widehat{\mathbf{C}}$ is regular and cocomplete, and the functors $F \widehat{\otimes} (-)$ are cocontinuous [45]. $\square$

However, the subunits in $\widehat{\mathbf{C}}$ are in general not well behaved.

Example 2.72. Consider the commutative monoid $M = [0,1) \times [0,\infty)$ under

$$(a,b) + (c,d) = \begin{cases} (a+c, b+d) & \text{if } a+c < 1 \\ (a+c-1, b+d+1) & \text{if } a+c \geq 1 \end{cases}$$

with unit $(0,0)$. Then M is a firm one-object category, but $\widehat{M}$ is not firm.

Proof. The identity $(0,0)$ represents the only subunit of the one-object category M, which is therefore firm. Appendix A proves that $\widehat{M}$ is not firm. $\square$

Moreover, $\widehat{\mathbf{C}}$ may have subunits that are not suprema of subunits of $\mathbf{C}$.

Remark 2.73. In general $\mathrm{ISub}(\widehat{\mathbf{C}})$ is not the free frame on $\mathrm{ISub}(\mathbf{C})$.

Proof. Consider a commutative unital quantale Q as a firm category. By their description in Appendix A, any subunit in $\widehat{Q}$ is given by a suitable downward closed subset $S \subseteq \downarrow e \subseteq Q$ such that $\forall x \in S\, \exists y,z \in S \colon x \leq yz$, and to be a subunit it suffices for S to be directed.

In particular, take $Q = [0,\infty]$ under the opposite order and addition. Then $\mathrm{ISub}(Q) = \{0,\infty\}$, whose free completion to a frame is its collection of downsets $\big\{\emptyset, \{\infty\}, \{0,\infty\}\big\}$. However, by the above description of subunits in $\widehat{Q}$ it is easy to see that

$$\mathrm{ISub}(\widehat{Q}) \supseteq \big\{\emptyset, \{\infty\}, [0,\infty], (0,\infty]\big\}$$

$\square$

Instead, to complete $\mathrm{ISub}(\mathbf{C})$ to a distributive lattice, preframe, or frame, we will consider certain full subcategories of $\widehat{\mathbf{C}}$.

Definition 2.74. A presheaf on a braided monoidal category $\mathbf{C}$ is *(finitely, directedly) broad* when it is naturally isomorphic to one of the form

$$\langle U, X\rangle \colon A \mapsto \{f \colon A \to X \mid f \text{ restricts to some } s \in U\}$$

for a (finitely bounded, directed) family U of subunits and an object X.

Write $\widehat{\mathbf{C}}_{\mathrm{brd}}$ ($\widehat{\mathbf{C}}_{\mathrm{fin}}$, $\widehat{\mathbf{C}}_{\mathrm{dir}}$) for the full subcategory of (finitely, directedly) broad presheaves. We will also write $\widehat{U}$ for $\langle U, I\rangle$, and $\widehat{X}$ for $\langle\{1\}, X\rangle$.

We will see below that the broad presheaves are precisely the colimits of the diagrams $D(\{\hat{s} \mid s \in U\}, \hat{X})$, and leave open the possibility of characterising when a given presheaf is broad in terms not referring to U or X.

The following lemma shows that broad presheaves are closed under (Day) tensor products and so form a monoidal category.

Lemma 2.75. *For any objects X, Y and families of subunits U, V in a stiff category $\mathbf{C}$, there is a (unique) natural isomorphism making*

$$\begin{tikzcd} \langle U,X\rangle \mathbin{\widehat{\otimes}} \langle V,Y\rangle \arrow[r, dashed, "\simeq"] \arrow[d, "u \mathbin{\widehat{\otimes}} v"'] & \langle U\otimes V, X\otimes Y\rangle \arrow[d, tail] \\ \widehat{X} \mathbin{\widehat{\otimes}} \widehat{Y} \arrow[r, "\simeq"] & \widehat{X\otimes Y} \end{tikzcd} \tag{2.8}$$

commute, where $U \otimes V = \{s \wedge t \mid s \in U, t \in V\}$*, and* u, v *are the inclusions.*

Proof. See Appendix A. □

We now describe the subunits in each completion.

Proposition 2.76. *If* $\mathbf{C}$ *is stiff, the subunits in* $\widehat{\mathbf{C}}_{\text{brd}}$ *(*$\widehat{\mathbf{C}}_{\text{fin}}$*,* $\widehat{\mathbf{C}}_{\text{dir}}$*) are the presheaves of the form* $\widehat{U}$ *for (finitely bounded, directed)* $U \subseteq \mathrm{ISub}(\mathbf{C})$.

Proof. Clearly $\widehat{U}$ is a subunit. Conversely, if $\eta \colon \langle U, X\rangle \to \widehat{I}$ is a subunit then

$$s_X = \eta_{S \otimes X}(s \otimes X) \colon S \otimes X \to I$$

will be proven to be a subunit in $\mathbf{C}$ for each $s \in U$.

Given this, let $U' = \{s_X \mid s \in U\}$, noting that $\widehat{U'}$ again belongs to each respective category, and consider the function $\langle U, X\rangle(A) \to \langle U', I\rangle(A)$ given by $((s \otimes X) \circ f) \mapsto s_X \circ f$. It is surjective by definition of U', clearly natural, and is well-defined and injective since

$$\begin{aligned} s_X \circ f = s'_X \circ f' &\iff \eta(s \otimes X) \circ f = \eta(s' \otimes X) \circ f' \\ &\iff \eta((s \otimes X) \circ f)) = \eta((s' \otimes X) \circ f') \\ &\iff (s \otimes X) \circ f = (s' \otimes X) \circ f' \end{aligned}$$

by naturality and injectivity of η.

Let us show that s_X is indeed a subunit. By stiffness of $\mathbf{C}$ each morphism $(s \otimes X)$ is monic, and so by the above argument s_X is, too.

Next we show $s_X \otimes S \otimes X$ is invertible. Notice that $\langle U, X\rangle = \langle \downarrow U, X\rangle$, so we may assume that U is idempotent. The fact that η is a subunit means precisely that each map

$$\langle U, X \otimes X\rangle(A) \to \langle U, X\rangle(A) \tag{$*$}$$

$$(s \otimes (X \otimes X)) \circ f \mapsto (s_X \otimes X) \circ f \tag{2.9}$$

is a well-defined bijection, where $f \colon A \to S \otimes X \otimes X$ and $s \in U$.

Now note that $S \otimes s_X \otimes X$ is monic, since by injectivity of $(*)$, $s_X \otimes X$ is monic, and it is easy to see from stiffness that for any subunit s and monic m that $S \otimes m$ is again monic. Moreover it is split epic and hence an isomorphism, since by surjectivity of $(*)$ there is some f with $(s_X \otimes X) \circ f = s \otimes X$, and $S \otimes (s \otimes X)$ is always split epic by idempotence of s. □

For any semilattice, as well as its downsets forming its free completion to a frame, recall that its free completion to a preframe is given by its collection of directed downsets [90, Theorem 9.1.5]; and that its free completion to a distributive lattice is given by its finitely bounded downsets [48, I.4.8], with (directed, finite) joins given by unions.

Corollary 2.77. *The subunits in* $\widehat{\mathbf{C}}_{\text{fin}}$, $\widehat{\mathbf{C}}_{\text{dir}}$, *and* $\widehat{\mathbf{C}}_{\text{brd}}$, *are the free completion of* $\mathrm{ISub}(\mathbf{C})$ *to a distributive lattice, preframe, and frame, respectively.*

Proof. For any $U, V \subseteq \mathrm{ISub}(\mathbf{C})$ it is easy to see that $\widehat{U} \leq \widehat{V} \iff U \leq {\downarrow}V$. In particular $\widehat{U} = \widehat{{\downarrow}U}$ as we have already noted. Hence by Proposition 2.76, subunits in each category correspond to the respective kinds of downset $U \subseteq \mathrm{ISub}(\mathbf{C})$. □

Next let us note that each of our constructions are again stiff.

Lemma 2.78. *If a monoidal category* $\mathbf{C}$ *is stiff, then so are* $\widehat{\mathbf{C}}_{\text{dir}}$, $\widehat{\mathbf{C}}_{\text{fin}}$ *and* $\widehat{\mathbf{C}}_{\text{brd}}$.

Proof. For any object $\langle U, X\rangle$ and subunit $V\colon \widehat{V} \to \widehat{I}$ in $\widehat{\mathbf{C}}_{\text{brd}}$ we need to show that the morphism $\langle U, X\rangle \otimes V$ is monic. This holds since the obvious morphism $\langle U, X\rangle \otimes \widehat{V} \to \widehat{X}$ factors over it, and is itself monic by equation (2.8) of Lemma 2.75.

By the same result, for the pullback property we must show each diagram

$$\begin{tikzcd}
\langle U \otimes V \otimes W, X\rangle \arrow[r, tail] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & \langle U \otimes W, X\rangle \arrow[d, tail] \\
\langle V \otimes W, X\rangle \arrow[r, tail] & \langle W, X\rangle
\end{tikzcd}$$

to be a pullback in $\widehat{\mathbf{C}}_{\text{brd}}$. For this it suffices to check that applying the diagram to each object A yields a pullback in **Set**, or equivalently that any morphism $f\colon A \to X$ factoring over $u \otimes w \otimes X$ and $v \otimes w' \otimes X$ for some $u \in U, v \in V$ and $w, w' \in W$ factors over $u' \otimes v' \otimes w'' \otimes X$ for some $u' \in U, v' \in V, w'' \in W$. But this follows easily from the pullbacks (2.4) taking $u' = u$, $v' = v$ and $w'' = w \wedge w'$, again for convenience assuming W to be idempotent. □

The next lemma shows that $\widehat{\mathbf{C}}_{\text{brd}}$ formally adds to $\mathbf{C}$ the colimits of the diagrams $D(U, X)$ for all suitable $U \subseteq \mathrm{ISub}(\mathbf{C})$ and objects X.

Lemma 2.79. *Let* $\mathbf{C}$ *be firm, and let* $U, V \subseteq \mathrm{ISub}(\mathbf{C})$ *be idempotent. Morphisms* $\alpha\colon \langle U, X\rangle \to \langle V, Y\rangle$ *of broad presheaves correspond to cocones* $c_s\colon S \otimes X \to Y$ *over* $D(U, X)$ *for which each* c_s *restricts to some* $t \in V$.

Proof. Given α and $s \in U$, by naturality we may define such a cocone by $c_s = \alpha_{S\otimes X}(s \otimes X)$. Conversely, given a cocone as above define

$$\alpha_A\big((s \otimes X) \circ g\big) = c_s \circ g$$

for each $g\colon A \to S \otimes X$. This is clearly natural and is well-defined; indeed if $(s \otimes X) \circ g = (t \otimes X) \circ h$ then since (2.4) is a pullback this morphism factors as $(s \otimes t \otimes X) \circ k$ for some k, then with $c_s \circ g = c_{s\wedge t} \circ k = c_t \circ h$ since the (c_s) form a cocone. Clearly these two assignments are inverses. □

Finally we can prove that our free constructions have the desired properties.

Theorem 2.80. *If* $\mathbf{C}$ *is a stiff category, then:*

- $\widehat{\mathbf{C}}_{\text{fin}}$ *has universal finite joins of subunits;*
- $\widehat{\mathbf{C}}_{\text{dir}}$ *has universal directed joins of subunits;*
- $\widehat{\mathbf{C}}_{\text{brd}}$ *is locale-based.*

Proof. Consider the final statement first. Lemma 2.78 makes $\widehat{\mathbf{C}}_{\text{brd}}$ stiff. Let $\mathcal{U}$ be an idempotent family of subunits in $\widehat{\mathbf{C}}_{\text{brd}}$. By Proposition 2.76, its elements are of the form $\widehat{U}$ for some $U \subseteq \mathrm{ISub}(\mathbf{C})$. Also, its supremum in $\mathrm{ISub}(\widehat{\mathbf{C}}_{\text{brd}})$ is given by $\langle \bigcup \mathcal{U}, I\rangle$ where we write $\bigcup \mathcal{U} = \bigcup\{U \mid \widehat{U} \in \mathcal{U}\}$.

Let $V \subseteq \mathrm{ISub}(\mathbf{C})$, and let Y be an object in $\mathbf{C}$. We have to prove that the inclusions $\widehat{U}\widehat{\otimes}\langle V,Y\rangle \to \bigcup\mathcal{U}\widehat{\otimes}\langle V,Y\rangle$ are a colimit of the diagram $D(\mathcal{U},\langle V,Y\rangle)$ in $\widehat{\mathbf{C}}_{\text{brd}}$. By Lemma 2.75, we may equivalently consider the inclusions

$$\langle U \otimes V, Y\rangle \hookrightarrow \langle (\bigcup \mathcal{U}) \otimes V, Y\rangle.$$

These certainly form a cocone. The questions is whether it is a universal one. Suppose that $\alpha_U\colon \langle U \otimes V, Y\rangle \to \langle W, Z\rangle$ is another cocone. Define a natural transformation $\beta\colon \langle (\bigcup \mathcal{U}) \otimes V, Y\rangle \to \langle W, Z\rangle$ by $\beta_A(f) = (\alpha_U)_A(f)$ for any $f\colon A \to X$ that restricts to $U \in \mathcal{U}$.

Now β is indeed well-defined, since if f also restricts to $U' \in \mathcal{U}$ then by the pullback (2.4), it also restricts to $U \cap U' \in \mathcal{U}$, so that $(\alpha_U)_A(f) = (\alpha_{U\cap U'})_A(f) = (\alpha_{U'})_A(f)$. By definition β is the unique natural transformation making the following triangle commute:

$$\begin{tikzcd}
\langle U \otimes V, Y\rangle \arrow[r, tail] \arrow[dr, "\alpha_U"'] & \langle (\bigcup \mathcal{U}) \otimes V, Y\rangle \arrow[d, "\beta"] \\
& \langle W, Z\rangle
\end{tikzcd}$$

Hence the inclusions indeed form a colimit, and $\widehat{\mathbf{C}}_{\mathrm{brd}}$ is locale-based. The proofs of the first two statements are identical, observing that if $U, V \subseteq \mathrm{ISub}(\mathbf{C})$ and $\mathcal{U} \subseteq \mathrm{ISub}(\widehat{\mathbf{C}}_{\mathrm{fin}})$ or $\mathrm{ISub}(\widehat{\mathbf{C}}_{\mathrm{dir}})$ are finitely bounded or directed, then so are $U \otimes V$ and $\bigcup \mathcal{U}$. □

We end this section by showing that the locale-based completion cannot be read in the traditional topological sense, in that broad presheaves are not sheaves for any Grothendieck topology.

Proposition 2.81. *There is a firm category* $\mathbf{C}$ *for which there is no Grothendieck topology* J *with* $\widehat{\mathbf{C}}_{\mathrm{brd}} \simeq \mathrm{Sh}(\mathbf{C}, J)$.

Proof. Suppose that $\widehat{\mathbf{C}}_{\mathrm{brd}}$ is a Grothendieck topos. Then it is a reflective subcategory of $\widehat{\mathbf{C}}$ [13, Proposition 3.5.4]. Hence $\widehat{\mathbf{C}}_{\mathrm{brd}}$ has a terminal object $\langle U, X \rangle$ that, because right adjoints preserve limits, must equal the terminal object of $\widehat{\mathbf{C}}$. Therefore, for all objects A of $\mathbf{C}$, the set $\langle U, X \rangle(A)$ must be a singleton. This means that for all objects A, there is a unique morphism $A \to X$ that restricts to some $s \in U$.

Suppose $\mathrm{ISub}(\mathbf{C}) = \{I\}$. Since every morphism restricts to I, now X must be a terminal object. But there exists a braided monoidal category $\mathbf{C}$ with only one subunit but no terminal object: any nontrivial abelian group. □

Remark 2.82. In future it would be natural to consider the above completions with presheaves valued in a category other than **Set** [14]. After all, Example 2.14 is enriched over complete lattices, Example 2.19 is enriched over abelian groups, and Example 2.24 is enriched over normed vector spaces. Proposition 2.71 holds for enriching categories $\mathbf{V}$ that are complete, cocomplete, locally small, and symmetric monoidal closed [45], covering all these examples. But an enriched version of Definition 2.74 would require taking the subobject of $[A, X]$ in $\mathbf{V}$ that restricts to some $s \in U$.

2.9 Universality of the completions

Finally, let us prove that the locale-based completion $\widehat{\mathbf{C}}_{\mathrm{brd}}$ and our other constructions $\widehat{\mathbf{C}}_{\mathrm{fin}}$ and $\widehat{\mathbf{C}}_{\mathrm{dir}}$ indeed have universal properties.

Definition 2.83. A *morphism* of categories with universal (finite, directed) joins of subunits is a braided monoidal functor $F\colon \mathbf{C} \to \mathbf{D}$ that preserves subunits and their (finite, directed) suprema. For short we call morphisms of categories with universal joins of subunits simply morphisms of locale-based categories.

Here, a functor F is monoidal when it comes equipped with coherent isomorphisms $\varphi_{A,B}\colon F(A)\otimes F(B) \to F(A\otimes B)$ and $\varphi\colon I \to F(I)$; these need to be invertible to make sense of preservation of subunits: if $s \in \mathrm{ISub}(\mathbf{C})$, then $\varphi^{-1}\circ F(s) \in \mathrm{ISub}(\mathbf{D})$.

By Lemma 2.68 and Theorem 2.70, a morphism is equivalently a braided monoidal functor $F\colon \mathbf{C} \to \mathbf{D}$ with $F\big(\operatorname{colim} D(U,X)\big) = \operatorname{colim} D\big(F(U),F(X)\big)$ for (finitely bounded, directed) idempotent $U \subseteq \mathrm{ISub}(\mathbf{C})$ and objects X of $\mathbf{C}$.

Definition 2.84. The *locale-based completion* of a braided monoidal category $\mathbf{C}$ is a monoidal functor $y\colon \mathbf{C} \to \mathbf{D}$ that preserves subunits such that $\mathbf{D}$ is locale-based, and any monoidal functor $\mathbf{C} \to \mathbf{E}$ into a locale-based category that preserves subunits factors as y followed by a morphism of locale-based categories G that is unique up to a unique monoidal natural isomorphism γ with $\gamma_y = 1_G$.

$$\begin{tikzcd}
\mathbf{C} \arrow[rr, "\substack{\text{monoidal,}\\ \text{preserves subunits}}"] \arrow[drr, "\substack{\text{monoidal,}\\ \text{preserves subunits}}"'] & & \mathbf{D} \arrow[d, dashed, "\text{locale-based}"] \\
& & \mathbf{E}
\end{tikzcd}$$

A *completion under universal finite* or *directed joins of subunits* of $\mathbf{C}$ is defined similarly.

Theorem 2.85. *If* $\mathbf{C}$ *is a stiff category, then via the Yoneda embedding its*

- *completion under universal finite joins of subunits is* $\widehat{\mathbf{C}}_{\mathrm{fin}}$*;*
- *completion under universal directed joins of subunits is* $\widehat{\mathbf{C}}_{\mathrm{dir}}$*;*
- *locale-based completion is* $\widehat{\mathbf{C}}_{\mathrm{brd}}$.

Proof. We prove the locale-based case, the others being identical. For any monoidal functor $F\colon \mathbf{C} \to \mathbf{D}$ into a locale-based category, we need to show that there is a morphism $\overline{F}\colon \widehat{\mathbf{C}}_{\mathrm{brd}} \to \mathbf{D}$ with $\overline{F} \circ y = F$, where y is the Yoneda embedding.

Because $\langle U, X\rangle = \langle \downarrow U, X\rangle$ for any $U \subseteq \mathrm{ISub}(\mathbf{C})$, we may assume that U is idempotent. Because F is monoidal, $F(U)$ is idempotent too. On objects, the requirement $\overline{F} \circ y = F$ forces us to define

$$
\begin{aligned}
\overline{F}\langle U, X\rangle &= \overline{F}\big(\operatorname{colim} D(y(U), \widehat{X})\big) \\
&= \operatorname{colim} D\big(\overline{F} \circ y(U), \overline{F} \circ y(X)\big) \\
&= \operatorname{colim} D\big(F(U), F(X)\big) \\
&\simeq \big(\bigvee F(U)\big) \otimes F(X).
\end{aligned}
$$

Now consider morphisms of (broad) presheaves. Any $\alpha\colon \langle U, X\rangle \to \langle V, Y\rangle$ induces a cocone $\alpha_s = \alpha_{S\otimes X}(s \otimes X)\colon\ S \otimes X \to Y$ over $D(U, X)$, where, as in Lemma 2.79, each such map factors through $t \otimes Y$ for some $t \in V$. Hence $F(\alpha_s)$ factors through $F(t) \otimes F(Y)$ and hence $\operatorname{colim} D\big(F(V), F(Y)\big) = \overline{F}\langle V, Y\rangle$, giving a morphism β_s as below.

$$
\begin{tikzcd}
F(S) \otimes F(X) \arrow[r, "\simeq"] \arrow[d, "\simeq"'] & F(S \otimes X) \arrow[r, "F(\alpha_s)"] & F(Y) \arrow[d, equal] \\
\overline{F}\langle \{s\}, X\rangle \arrow[rr, "\overline{F}(\alpha_s \circ (-))"] \arrow[d] \arrow[rrd, dashed, "\beta_s"] & & \overline{F}(\widehat{Y}) \\
\overline{F}\langle U, X\rangle \arrow[rr, dashed, "\overline{F}(\alpha)"'] & & \overline{F}\langle V, Y\rangle \arrow[u]
\end{tikzcd}
$$

By Lemma 2.68, the upper row forms a cocone over $D\big(F(U), F(X)\big)$ with s ranging over U. Because the vertical composite on the right is monic, the β_s also form a cocone (after composition with the upper left vertical isomorphism). But $\overline{F}\langle U, X\rangle$ is a colimit, so there is a mediating map $\overline{F}(\alpha)$ making the diagram commute. Uniqueness of this map makes $\overline{F}$ functorial. Given our definition of $\overline{F}$ on objects, this assignment $\overline{F}(\alpha)$ is unique with $\overline{F} \circ y = F$, since for each $s \in S$ the lower square commutes by functoriality, with the lower left vertical morphisms forming a colimit.

Next, $\overline{F}$ may readily be checked to be (strong) braided monoidal:

$$\begin{aligned}\overline{F}(\langle U,X\rangle \widehat{\otimes} \langle V,Y\rangle) &\simeq \overline{F}\langle U\otimes V, X\otimes Y\rangle \\ &\simeq \big(\bigvee_{s\in U, t\in V} F(s)\wedge F(t)\big)\otimes F(X)\otimes F(Y) \\ &\simeq \bigvee F(U)\otimes \bigvee F(V)\otimes F(X)\otimes F(Y) \\ &\simeq \overline{F}\langle U,X\rangle \otimes \overline{F}\langle V,Y\rangle\end{aligned}$$

By construction $\overline{F}$ preserves subunits because $\overline{F}\langle U,I\rangle = \bigvee F(U)$, as well as their suprema:

$$\overline{F}\big(\bigvee_{U\in\mathcal{U}}\langle U,I\rangle\big) = \overline{F}\langle \bigcup \mathcal{U}, I\rangle \simeq \bigvee_{U\in\mathcal{U}}\bigvee_{s\in U} F(s) \simeq \bigvee_{U\in\mathcal{U}} \overline{F}\langle U,I\rangle$$

Hence $\overline{F}$ is indeed a morphism of locale-based categories.

Finally, we must show for any other morphism $\overline{F}'$ with $\overline{F}'\circ y = F$ that there is a unique monoidal natural isomorphism $\gamma\colon \overline{F}\to\overline{F'}$ with $\gamma_y = 1_F$. But this follows from the uniqueness of $\operatorname{colim} D\big(F(U),F(X)\big)$ up to unique isomorphism, and our statement above on the uniqueness of $\overline{F}(\alpha)$. □

We leave open the question how these completions relate to the free cocompletions in a left exact context in the case of toposes [68].

Each construction is functorial; we consider the locale-based case in detail. Write **LocBased** for the category of locale-based categories and their morphisms, and **Stiff** for the category of stiff categories and braided monoidal functors that preserve subunits.

Proposition 2.86. *The map* $\mathbf{C}\mapsto\widehat{\mathbf{C}}_{\mathrm{brd}}$ *defines a functor* **Stiff** $\to$ **LocBased**.

Proof. For any $F\colon \mathbf{C}\to\mathbf{D}$ in **Stiff**, define $\widehat{\mathbf{C}}_{\mathrm{brd}}\to\widehat{\mathbf{D}}_{\mathrm{brd}}$ on objects by $\langle U,X\rangle\mapsto\langle F(U),F(X)\rangle$. We have seen that it suffices to consider when U is idempotent. By Lemma 2.79, morphisms $\alpha\colon\langle U,X\rangle\to\langle V,Y\rangle$ are equivalently cocones over $D(U,X)$ each of whose legs factors over $t\otimes Y$ for some $t\in V$. Map such a cocone c_s to the cocone $F(c_s)$ over $D(F(U),F(X))$. This is well-defined by Lemma 2.68, and clearly functorial. □

It follows from Theorem 2.85 that the locale-based completion functor of the previous proposition is a left *biadjoint* to the forgetful functor **LocBased** $\to$ **Stiff**, when we make each category a strict 2-category with 2-cells being monoidal natural transformations (for this it suffices to check that each Yoneda embedding $\mathbf{C}\to\widehat{\mathbf{C}}_{\mathrm{brd}}$ is a *biuniversal arrow* [28, Theorem 9.16]).

The other constructions $\mathbf{C} \mapsto \widehat{\mathbf{C}}_{\mathrm{fin}}$ and $\mathbf{C} \mapsto \widehat{\mathbf{C}}_{\mathrm{dir}}$ similarly give left biadjoints; write **UnivFin** or **UnivDir** for the category of categories with universal finite or directed joins.

Theorem 2.87. *The following cube of forgetful functors commutes, all functors in the top face have left biadjoints, and the rest have left adjoints.*

$$\begin{tikzcd}
& \mathbf{UnivDir} \arrow[dl] \arrow[dd] & & \mathbf{LocBased} \arrow[ll] \arrow[dl] \arrow[dd, "\mathrm{ISub}"] \\
\mathbf{Stiff} \arrow[dd] & & \mathbf{UnivFin} \arrow[ll] \arrow[dd] & \\
& \mathbf{PreFrame} \arrow[dl] & & \mathbf{Frame} \arrow[ll] \arrow[dl] \\
\mathbf{SemiLat} & & \mathbf{DistrLat} \arrow[ll] &
\end{tikzcd}$$

Proof. All functors in the bottom face have a left adjoint [49, Lemma C1.1.3]. Explicitly: the free frame on a preframe is given by taking its Scott closed subsets [9, Proposition 1], and we have already mentioned the free frame, preframe or distributive lattice on a semilattice. Observe that all these free constructions take certain types of downward closed subsets. Therefore they can be categorified from posets to categories that have universal joins of these types of subsets of subunits. The universal property of Theorem 2.85 then holds in each case. Hence all functors in the top face of the cube have a left biadjoint. Finally, all vertical functors have a left adjoint as in Example 2.11. $\square$

Chapter 3

Protocols

The goal of this chapter is to demonstrate how the theory presented in Chapter 2 can be used in practice. Now that we have a categorical way to speak about space and time, we can revisit how protocols are categorically modelled with the goal of capturing more information.

First, we define a diagrammatic category that allows us to focus on the causal structure of a protocol. This is simple but it is not clear how to upgrade it into a model that provides a more detailed account of the regions involved. For this purpose, we have a second construction, this time phrased in terms of subunits, which can be seen as a generalisation of the first one. In a sense that will be made precise below, this formalises the idea of having morphisms of a category being supported in subunits of another one. Thanks to this, we can have a category modelling 'what' can happen, while another category can be dedicated to 'where' and 'when' things can take place.

3.1 Background: logical clocks

Recall that *distributed system* is loosely defined as a collection of agents or computers running in parallel, each capable of taking independent actions as well as of exchanging messages with the other agents (see [1, 60] for more on this topic).

As argued in [60, Chapter 3], on which this brief introduction is mostly based, causal reasoning is very prominent and useful when analysing and designing distributed algorithms. Naively, one way to capture the potential for causal interaction between the different events in a distributed computation would be to keep track of the coordinates, both for time and space, at which they happen. However, coordinates depend on the frame of reference, and their use can entail problems of synchronisation and precision which would be very hard to avoid in practice. Fortunately, it is not necessary to refer to physical coordinates when capturing causal structure, because so-called logical clocks are enough.

Remark 3.1. For the following definitions, we assume that each event in the domain of the logical clock function belongs to one of a number of processes running in parallel, and that these processes can exchange messages, the sending and receiving of which constitute events. Each event carries its timestamp with it, and messages carry the timestamp of the event that sends them. So, the time of an event may be determined by only looking at those events right before it. Graphically, we will draw such sets of events as follows, where the horizontal lines correspond to the different local processes, the dots in them to events, and the arrows between them to messages. Time runs from left to right.

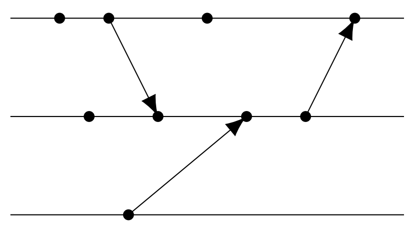


Figure 3.1: a protocol involving three processors

Definition 3.2. A *logical clock* C is a function from a set of events in a distributed system to a poset $(T, \trianglelefteq)$, called the *time domain*, with $e_1 \preceq e_2 \Rightarrow C(e_1) \trianglelefteq C(e_2)$ for any two events e_1, e_2. In this section, $\preceq$ is the *happens-before* relation: $e_1 \preceq e_2$ whenever there is a path from e_1 to e_2 in the direction of increasing time, meaning that e_1 has the potential to influence e_2. We call $C(e)$ the *time* of, or the *timestamp* of, event e.

Example 3.3. *Scalar time* or *Lamport timestamps* is a logical clock C_L with the non-negative integers as time domain, defined by the following rules.

1. When an event receives a message m with time $C_L(m)$, then its timestamp is $\max(C_L(e), C_L(m)) + 1$, where e is the previous event in the same processor. If there is no previous event in the same processor, simply assign time $C_L(m) + 1$.

2. For any other event, increase by one the time of the previous event in the same processor, or assign 1 if it is the first in that processor.

We can imagine that there is a clock at every process which keeps track of how many events precede a given one. From this point of view, this logical time represents

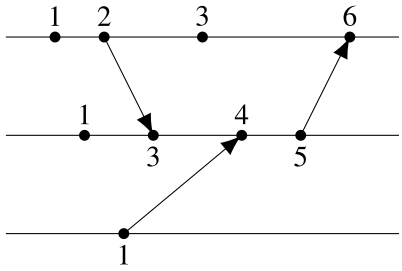


Figure 3.2: the protocol in Figure 3.1, now with Lamport timestamps

the local knowledge of how many events are taking place. This results in a valid logical clock, because $a \preceq b \Rightarrow C_L(a) \trianglelefteq C_L(b)$. Note, however, that the opposite is not true in general. One way to get a stronger correspondence between logical clocks and causal structure is via vector clocks.

Instead of keeping track only of the local knowledge of one process, each local clock can also record the knowledge of all other clocks, by using vectors instead of integers as the time domain. In this case, it can be shown that $a \preceq b \iff C_V(a) \trianglelefteq C_V(b)$.

Definition 3.4. *Vector clocks* is a logical clock C_V with n-dimensional vectors as time domain, where n is the number of processes in which the events can be. Write the time of an event e as $C_V(e) = \{C_V(e)_i\}_{i=1,\dots,n}$, and define C_V via the following rules.

1. When an event in process j receives a message with time $\{C_V(m)_i\}_{i=1,\dots,n}$, assign to it time

   $$\{\max(C_V(m)_1, C_V(e)_1), \dots, \max(C_V(m)_j, C_V(e)_j)+1, \dots, \max(C_V(m)_n, C_V(e)_n)\},$$

   where e is the previous event in process j (if it does not exist, $C_V(e)$ is the zero vector).

2. For any other event a in process j, simply increase by one the j-th component of the time of the previous event in the same processor. If there is no previous event, assign time $C_V(a)_j = 1$, $C_V(a)_i = 0 \quad \forall i \neq j$.

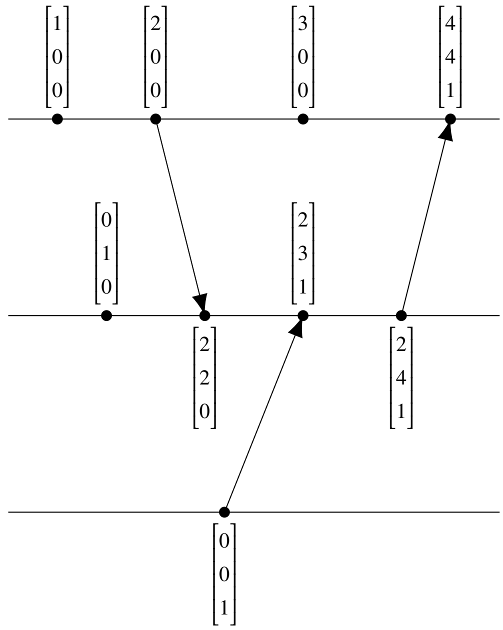


Figure 3.3: the same protocol as in Figures 3.1 and 3.2, now timestamped with vector clocks

3.2 Diagrams

In the standard categorical approach to physics, the morphisms of a category **C** model exactly which processes are allowed to happen. In this section, we combine this idea with the diagrammatic representation of distributed protocols as in Figure 3.1, by labelling the vertices of the diagram (that is, the events) with morphisms of **C**.

Definition 3.5. Given a category **C**, we define the monoidal category $\mathbf{Diag}(\mathbf{C})$ of *diagrams*[1] *over* **C**. The objects in $\mathbf{Diag}(\mathbf{C})$ are the non-zero natural numbers. The set of morphisms $\mathbf{Diag}(\mathbf{C})(n,m)$ is empty whenever $n \neq m$. A morphism or *diagram* of type $n \to n$ is a directed graph whose vertices are partitioned into n components, and where each vertex is labelled by a morphism of **C**. Composition is graphical: simply draw one graph after the other and draw edges from the latest vertices of the first to the earliest of the second. Similarly, by drawing the graphs on top of each other we get the tensor product, as exemplified in Figure 3.4.

An object indicates the number of parties involved, and for simplicity we only consider protocols with the same number of initial and final parties. The partition of the vertices captures how the events are distributed among the parties.

[1]In category theory, *diagram* can be used to refer to a functor [63, Definition 5.1.18]. Here, we use the term in the sense of diagrammatic or graphical calculus [81].

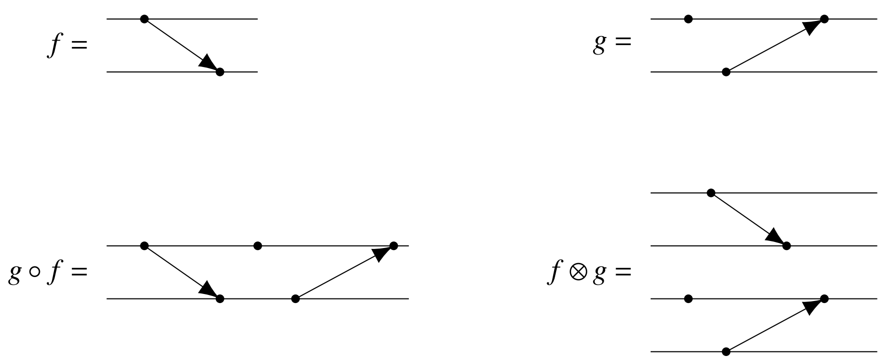


Figure 3.4: examples of morphisms

This way to graphically define composition and tensor product is the same as in the graphical calculus of CQM [81].

We interpret the message arrows as indicating that, before the receiver can start their process, they need to wait for the sender's to finish. This may be simply because a resource that the receiver needs is being used by the sender, in which case the message can represent transfer of the resource or a notification that the resource is now free. More generally, it may be because the message contains some other information that is crucial to the receiver. This point of view on messages allows us to easily characterise deadlock situations, in which different parties in a distributed computation are waiting for each other so no progress can be made.

Definition 3.6. We say that a diagram in $\mathbf{Diag}(\mathbf{C})$ *features deadlock* if and only if it contains a cycle.

Diagrams are well-suited for protocols that adhere to the assumptions in Remark 3.1, and can easily capture causal structure through the position of the events within a party and the messages between parties. Further, as we will see in Section 3.5, if we choose a convention to interpret the parties as labels for locations, we can use diagrams to capture some spacetime information on the protocol.

However, diagrams are not a natural setting for those protocols that do not have a clear partition into a finite number of parties. For instance, protocols will in general involve many different spacetime regions which will not be easily separable into a finite number of parties and messages between them. Conversely, there is no systematic way to interpret the parties as labels for locations, which would be necessary to add information on the locations of events.

Therefore, we want to introduce a model for protocols that do not necessarily adhere to the assumptions in Remark 3.1, and in which different kinds of spacetime information, including toy models and more realistic ones, can be included thanks to the use of subunits. This will be the goal of the next section.

3.3 The category of protocols

As we explained in the introduction, in categorical semantics it is common to interpret the objects of a category $\mathbf{C}$ as systems, and the morphisms as processes that transform a system into another. However, if we model a process as a morphism, we are in general forgetting all details of 'where' in spacetime it takes place. We can use subunits as a way to remedy this. Indeed, if the subunits of a category $\mathbf{D}$ give us a model of spacetime, we can then see events as pairs (f, s) where f is a morphism in $\mathbf{C}$ and $s \in \mathbf{D}$ informs of its support. Then, we can regard a protocol as a collection of events over some spacetime $\mathbf{D}$, together with a causal structure.

Note that we could, in principle, have $\mathbf{C} = \mathbf{D}$, but in practice it is very difficult to have, in a single category, the desired morphisms as well as the subunits that give a model of the relevant spacetime.

Definition 3.7. Let $\mathbf{C}$ be braided monoidal and $\mathbf{D}$ firm. Then, we define the category $\mathbf{Prot}(\mathbf{C}, \mathbf{D})$ *of protocols in* $\mathbf{C}$ *over* $\mathbf{D}$. An object A is a finite (possibly empty) set $A_U = \{(f_i, s_i)\}_{i=1,\dots,N}$ with f_i a morphism in $\mathbf{C}$ and $s_i \in \mathrm{ISub}(\mathbf{D})$ for all i; together with a *happens-before* preorder $\sqsubseteq_A$ on A_U. We call the elements of A_U *events*. We will assume protocols to be *thin*, in the sense that they contain at most one event (f, s) per subunit s. Call the set of subunits involved in a protocol A its *network* , $\mathcal{N}(A) := \{s \mid (f, s) \in A_U\}$. Finally, the morphisms are as follows.

$$\mathbf{Prot}(\mathbf{C}, \mathbf{D})(A, B) = \begin{cases} \{*\}, & \text{if } \mathcal{N}(A) \subseteq \mathcal{N}(B) \\ \emptyset, & \text{otherwise} \end{cases} \tag{3.1}$$

We can see the category of protocols as a way to formalise the idea of taking the usual CQM formulation and adding locations to (some of) the morphisms that compose to a protocol. An early idea on how to do this was discussed in [26]. This was based on restricting each of the morphisms f that form a protocol to the locations where they happen, via the restriction functor $f \mapsto f \otimes 1_S$. This was an interesting proof of concept that allowed us to check that the notions we had defined (subunits, support, restriction and causal structure) capture the intuitions we wanted, but it only makes sense when assuming that the category $\mathbf{C}$ contains all the information on what states, processes and locations are possible (respectively: objects, morphisms and subunits). In general, it is difficult to have, in a single category, the desired objects and morphisms as well as the desired subunits. Here, we follow a similar idea while allowing for processes and locations to be represented by different categories.

Happens-before tells us about the causal structure of the protocol, just like messages and the position within a party do in the diagrammatic setting above.

Note that the condition of thinness does not limit the amount of processes that we can specify as happening in one location. For instance, if we want to describe that both f and g are happening in s, then we can simply include the event $(f \otimes g, s)$ in the protocol.

Two protocols are isomorphic if and only if their networks are equal. In that case, they are canonically isomorphic, in the sense that there is only one morphism between them.

Because of thinness, there is a one to one correspondence between preorders $\sqsubseteq_A$ on a set of events A and preorders $\sqsubseteq_{\mathcal{N}(A)}$ on its network, and morphisms can equivalently seen as preorder maps between the networks.

Proposition 3.8. *Call $\mathcal{N}(\mathbf{D})$ the semilattice $(\mathcal{P}(\mathrm{ISub}(\mathbf{D})), \subseteq, \cap)$ seen as a monoidal category. Then, there is a forgetful functor $U\colon \mathbf{Prot}(\mathbf{C}, \mathbf{D}) \to \mathcal{N}(\mathbf{D})$ given by $A \mapsto \mathcal{N}(A)$, and this is monoidally right adjoint to the free functor $X \mapsto \{(1_I, s) \mid s \in X\}$ together with $s \sqsubseteq t \iff s \leq t$ in $\mathrm{ISub}(\mathbf{D})$.*

Proof. Because all isomorphisms are canonical, we just need to check that the networks are the same. □

In what follows, we describe a monoidal structure for $\mathbf{Prot}(\mathbf{C}, \mathbf{D})$ that gives us the same subunits as $\mathbf{D}$, as well as meaningful notions of support and restriction. As a result, we can understand this category as allowing us to formally speak about morphisms in $\mathbf{C}$ being supported in subunits of $\mathbf{D}$.

Definition 3.9. Given two protocols A and B, set

$$(A \otimes B)_U = \{(f \otimes g, s) \mid (f, s) \in A, (g, s) \in B\}$$

together with $s \sqsubseteq_{\mathcal{N}(A \otimes B)} t \iff s \sqsubseteq_{\mathcal{N}(A)} t$ and $s \sqsubseteq_{\mathcal{N}(B)} t$. For morphisms, if we have $\alpha : A \to A'$ and $\beta : B \to B'$, then it is clear that there is a unique morphism $A \otimes B \to A' \otimes B'$.

It is easy to see that we get a monoidal structure with $I = \{(1_{I_{\mathbf{C}}}, s) \mid s \in \mathrm{ISub}(\mathbf{D})\}$ as the tensor unit, together with $s \sqsubseteq_{\mathcal{N}(I)} t$ if and only if $s \leq t$ as subunits, and with canonical braiding, associator and unitors.

With this choice of monoidal structure, the category is determined, up to monoidal equivalence, by the spacetime.

Proposition 3.10. *There is a monoidal equivalence of categories between* $\mathbf{Prot}(\mathbf{C}, \mathbf{D})$ *and* $\mathbf{Prot}(\mathbf{C}', \mathbf{D})$*, for every* $\mathbf{C}$ *and* $\mathbf{C}'$*.*

Proof. This is witnessed by functors F that map protocol A to $\{(1_I, s) \mid s \in \mathcal{N}(A)\}$ together with $s \sqsubseteq_{FA} t \iff s \sqsubseteq_A t$. □

Proposition 3.11. $\mathrm{ISub}(\mathbf{Prot}(\mathbf{C}, \mathbf{D})) = \mathcal{P}(\mathrm{ISub}(\mathbf{D}))$ *as frames, and* $\mathbf{Prot}(\mathbf{C}, \mathbf{D})$ *is locale-based.*

Proof. Every protocol A can be seen as a subunit, together with the canonical embedding into the tensor unit. It is easy to see that $A \leq B$ as subunits if and only if $\mathcal{N}(A) \subseteq \mathcal{N}(B)$. Therefore, two subunits are equivalent if and only if they have the same networks. As a result, every subunit has a canonical representative consisting only of events of the form $(1_{I_{\mathbf{C}}}, s)$ for some s. So, there is a one-to-one correspondence between subunits of $\mathbf{Prot}(\mathbf{C}, \mathbf{D})$ and sets of subunits of $\mathbf{D}$. The straightforward nature of the morphisms makes it easy to check that the category is locale-based. □

Proposition 3.12. *A protocol A has support in B if and only if $\mathcal{N}(A) \subseteq \mathcal{N}(B)$, and the restriction* $\mathbf{Prot}(\mathbf{C}, \mathbf{D})|_B$ *is the full subcategory of protocols A supported in B.*

Proof. Follows from the definition of support. □

Because the category is locale-based, we can speak about 'the' support of a protocol, that is, the smallest region where it is supported. This corresponds exactly to its network, or in other words the set of locations that appear in its events.

As in the category of diagrams, it is easy to characterise deadlock. Furthermore, as we will see in Definition 3.31, we can characterise exactly when the causal structure of a protocol is in accordance with the possibilities that the spacetime offers.

Definition 3.13. We say that a protocol *features deadlock* whenever there are two distinct events (f, s), (g, t) such that $(f, s) \sqsubseteq_A (g, t)$ and $(g, t) \sqsubseteq_A (f, s)$.

Note that this does not imply that the two events that witness the deadlock have to be equal, since happens-before is just a preorder. Featuring deadlock is different from a protocol having the potential for deadlock, which would be the case whenever there are two events (f,s), (g,t) that can causally influence each other. We will formalise the possibility of causal interaction, as opposed to the necessity implied by the happens-before preorder, when we introduce causal structures in Section 3.6.

3.4 Categories of timestamps

Toy model of spacetime

Including some information on the locaions of events does not necessarily entail a realistic set of coordinates or times. For instance, it is simpler to consider labels for locations and times, since then we do not have to worry about synchronising the clocks of different observers. The theory of subunits allows us to choose any semilattice as our notion of 'space and time', so we can in particular treat toy models and realistic ones in the same categorical footing.

As explained above, logical clocks assign, to each event in a distributed protocol, a timestamp that belongs to a certain poset called the time domain. Since subunits form a semilattice and, in particular, a poset, we can take them to act as a kind of time domain. With this motivation in mind, let us define a family of semilattices (sets of subunits, when regarded categorically) whose elements allow us to label not only different times but also different locations.

Definition 3.14. The semilattice $\mathbf{Clock^n}$ of *clocks* consists of all n-tuples of integers $(x_1,\dots,x_n)$ and a top element $\top$, with partial order

$$(x_1,\dots,x_n) < (y_1,\dots,y_n) \iff \forall k : x_k \leq y_k, \quad \exists k' : x_{k'} < y_{k'}$$
$$(x_1,\dots,x_n) < \top$$

and with meet

$$(x_1,\dots,x_n) \wedge (y_1,\dots,y_n) = (\min(x_1,y_1),\dots,\min(x_n,y_n)),$$

The semilattice $\mathbf{Loc^m}$ of *locations* consists of the subsets of $\{1,\dots,m\}$, which form a semilattice with the partial order

$$x \leq y \iff y \subseteq x$$

and union as the meet.

Finally, the category $\mathbf{TStamp^{n,m}}$ of *timestamps* is the result of viewing the product semilattice $\mathbf{Clock^n} \times \mathbf{Loc^m}$ as a monoidal category. Write the elements as $(x_1, \ldots, x_m)_j$, where $j \subseteq \{1, \ldots, m\}$ is the location. Write $\mathbf{TStamp^n}$ for $\mathbf{TStamp^{n,n}}$.

As we wanted, we get a category whose subunits are timestamps. In the examples of Section 3.5, we will see how the clocks and locations do not necessarily have to model time or position, but can instead be used for other purposes.

Allowing for $n \neq m$ in the category of timestamps gives more flexibility when modelling protocols. For instance, and as we will see below, for quantum circuits it is enough to have $n = 1$ and m be the number of qubits. Instead, if we want to be able to reconstruct a protocol presented diagrammatically (as in Figure 3.1) from the timestamps of its events alone, we require n and m to both equal the number of parties involved.

Realistic model of spacetime

We can also have a semilattice whose elements are points or regions of a given manifold, as a more realistic model of spacetime. Recall that by a spacetime manifold $\mathcal{M}$ we mean Lorentzian manifold X with a time orientation [72], since this is the model commonly used in physics.

Given $\mathcal{M}$, a natural choice from the theoretical point of view would be to consider the lattice of open subsets. However, in this chapter we are motivated by allowing flexibility when modelling events and protocols occurring in spacetime, which in principle do not have to correspond to regions with special mathematical properties. Therefore, given a spacetime manifold $\mathcal{M}$, we will consider the locale $\mathcal{P}(\mathcal{M})$ of arbitrary regions, together with $\subseteq$ as the order. It follows that, viewed as a monoidal category, this locale has the intersection $\cap$ as tensor produt, and that its subunits are the arbitrary regions of spacetime, as wanted.

3.5 Examples with toy model of space and time

3.5.1 Diagrams, categorically

Recall that, in the category $\mathbf{Diag}(\mathbf{C})$ of diagrams over $\mathbf{C}$, a morphism is said to feature deadlock when its graph has any cycle. Given a diagram without deadlock in $\mathbf{Diag}(\mathbf{C})(n,n)$, which we interpret as a protocol involving n parties or locations, we can view it as an object A in $\mathbf{Prot}(\mathbf{C}, \mathbf{TStamp^{1,n}})$. Note that this spacetime corresponds to having Lamport clocks as logical time, but the following presentation would work with vector clocks or any other logical clocks as well, since all we need is to have a good happens-before relation and to keep track of locations in the timestamps.

Let us describe the construction by which, given a diagram, we can get an object in the category of protocols. For each vertex f in the diagram, we want to have an event $(f, C_L(f)_j) \in A$ where j is the location of the vertex, that is, the party to which it belongs. For the happens-before preorder on A, let us first recall the definition of covering relation.

Definition 3.15. Given a preordered set $(X, \leq)$, we can define a new relation as follows.

$$x \lessdot y \iff x \leq y, \nexists z \colon x \leq z \leq y$$

In this case, we say that y covers x. In our context, we will say that an event (g, t) *happens right after* (f, s) if the former covers the latter with respect to the happens-before preorder.

To define the happens-before preorder for A, we want a vertex g to be right after another f if and only if, either they are in the same location and g appears right after in the diagram, or there is a message from f to g. Then, happens-before is just the transitive closure of the relation we get in this way.

Conversely, given a protocol as an object $A \in \mathbf{Prot}(\mathbf{C}, \mathbf{TStamp^{1,n}})$, we can recover without ambiguity the diagram that would induce it. We simply need, for every event in A, a vertex placed in the location described by the timestmap. Within each location, events are totally ordered according to their logical time. To fill in the messages into the diagram, inspect the covering relation for the happens-before preorder of A.

It is clear that these translations, from diagrams to objects and back, are inverse to each other. Thus, we have the following.

Lemma 3.16. $\mathbf{Diag}(\mathbf{C})(n,n) \simeq \mathbf{Prot}(\mathbf{C}, \mathbf{TStamp^{1,n}})$ *for every* $n \in \mathbb{N}$.

This roughly means that, when we focus on toy models of space and time, diagrams and objects in the category of protocols are equivalently expressive, as will be exemplified in Section 3.5.

Capturing the content of messages

In general, we think of an event (f,s) in a protocol $A \in \mathbf{Prot}(\mathbf{C},\mathbf{D})$ as saying that a process f takes place within the region s. Similarly to how we can have labels for times and places, we can also have labels that denote whether an event is a message, and between which parties. One way to do this is to consider, when n parties are involved in the protocol, a total of n^2 locations, with the following convention: the labels $\{1,\dots,n\}$ denote the parties involved, and $\{n+1,\dots,n^2\}$ denote that the event is a message and inform of which parties are sending and receiving it. Write $\langle k,l\rangle$ to mean that the event is a sent by the k-th party and received by the l-th, and simply order these lexicographically.

Label	**Meaning**
1	1st party
2	2nd party
3	3rd party
4	$\langle 1,2\rangle$
5	$\langle 1,3\rangle$
6	$\langle 2,1\rangle$
7	$\langle 2,3\rangle$
8	$\langle 3,1\rangle$
9	$\langle 3,2\rangle$

Table 3.1: example of the convention with three parties.

Thanks to this, we can have a different intepretation for events (f,s) which are labelled as messages. For instance, we can view f as the content that the message is carrying. As we mentioned previously, messages may be simply a way to symbolise that a resource which the receiver needs has now been freed by the sender. If the resource or state is modelled by object $A \in \mathbf{C}$, then the event $(1_A, s)$, when labelled as a message, can be used to mean that the resource being freed and sent is A.

Remark 3.17. For the rest of the section, we will consider diagrams that follow the assumptions laid out in Remark 3.1, except that messages are allowed to carry their own timestamps, which may differ from their sender's.

Given a diagram of this kind with n parties, let us describe a way to assign timestamps to both the events and the messages. First, draw a new diagram with n^2 parties and place the messages as events in the appropriate party according to the convention above. For every message, draw an arrow from the sending event to the message, and from the message to the receiving event. Finally, assign timestamp $(x_1,\dots,x_{n^2})_j$ to each event and message, where $(x_1,\dots,x_{n^2})$ is the vector time of the event and j is the party or set of parties to which it belongs.

The reason for redrawing the given diagram is to transform it, without loss of information, into one where we can apply the standard vector clocks, since these are not defined for diagrams where the messages carry their own timestamps.

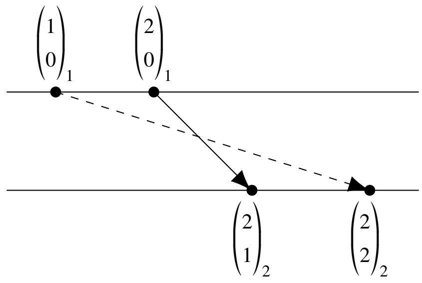
Note that this timestamping method does not quite define a new logical clock, since these timestamps also keep track of the locations. Without locations, however, we would get a logical clock which is just vector clocks on the auxiliary diagram. On the positive side, by keeping track of locations we have a bijection between diagrams and the collections of timestamps we obtain this way.

Proposition 3.18. *There is a bijection between diagrams adhering to the assumptions in Remark 3.17 and the collections of timestamps they induce.*

Proof. Given a collection of timestamps of the form $(x_1,\dots,x_{n^2})_j$ that was induced by a diagram via this construction, we can fully recover the diagram. First, draw n locations. Focusing on one location at the time, place in increasing order all the events that happen in that location. To place the messages, derive which are the sending and receiving locations for each message by looking at the location type, and finally derive which particular events within those locations are sending and receiving by inspecting the vector clocks times. □

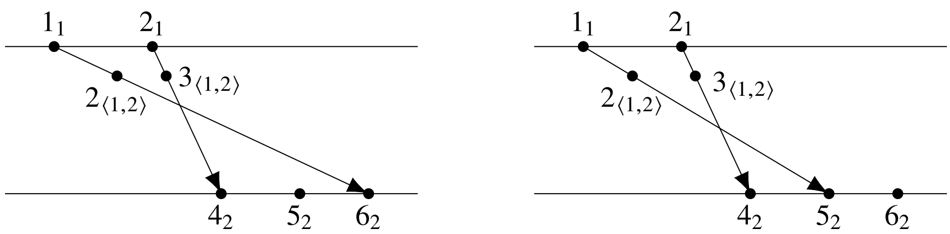
We need to keep track of locations if we want to avoid ambiguity when reconstructing the diagram from just its collection of timestamps. Furthermore, it would not be enough to consider regular vector clocks together with location information, as the next example demonstrates.

Example 3.19. Consider the following diagram, which has been timestamped using regular vector clocks with locations. In it, the timestamps would stay the same whether or not the dashed message is included, since we are not allowing messages to carry their own timestamps.



Using less components to keep track of time (for instance using Lamport timestamps instead of vector clocks) can also result in ambiguity, even if we allow messages to carry their own timestamps, as the following example demonstrates.

Example 3.20. These two diagrams share the same timestamps, if time is just an integer, even if we allow messages to have their own timestamps.



In general, when we consider an arbitrary protocol $A \in \mathbf{Prot}(\mathbf{C}, \mathbf{D})$, messages and events are both modelled in the category of protocols as events (f, s), and we have no way to distinguish them. Therefore, events (f, s) are forced into the interpretation of f as a process that takes place, and not as the content of a message. This happens, for instance, in the case of realistic locations, as we will see below.

We can see this as a trade-off: on the one hand, we can have subunits be labels, which lets us have different interpretations for events; and on the other hand we can have subunits capturing more detailed or realistic spacetime information, in which case we are forced to have a single interpretation for all events. The use of diagrams to model the contents of messages, as well as this trade-off, will be exemplified in Section 3.7.

3.5.2 Music

As a simple concrete example, let us consider music scores as protocols, and more concretely as objects in $\mathbf{Prot}(\mathbb{N}, \mathbf{TStamp^{1,14}})$. Each event will be an instruction for a note or set of notes to become active (that is, to start making sound) or inactive for a specified duration, and the happens-before relation will inform of when should the instruction be executed, as well as of the ordering of the notes.

We use the natural numbers, seen as a posetal category, to encode the duration of the notes. As a convention, we can choose number 1 to refer to the duration of a whole note, and an arbitrary number n to refer to a duration $1/n$. When a music score has notes longer than a whole note, it can always be rewritten to an equivalent one whose longest note is a whole note, and we will assume music is written in this form. Alternatively, we could have chosen number 1 to mean the duration of the whole piece, so n would represent a duration of $1/n$ that of the whole piece. But, by having 1 represent a whole note, we make encoding and decoding music as a protocol much easier and closer to the traditional musical notation, since for instance number 4 then represents a quarter note. Another advantage of this convention is that it is easier to combine pieces of music one after the other to create longer ones[2].

To allow for polyphony in which different notes can start and stop at different times, we will need silences to be local, or in other words each note needs labels for both the instructions of note activation and silence. We will treat these instructions as different locations, so since there are 7 notes (not counting sharps and flats, for this example) we end up with 14 labels. Let us denote these labels as $\{A, B, \ldots, G, A^*, B^*, \ldots, G^*\}$, where A means that the note activates and A^* is for silence.

Naively, it may seem that having the timestamps specify both the note and the duration could be enough as the spacetime for music. More precisely, that would mean taking the spacetime to be $\mathbf{Loc^{14}}$. However, in that case, thinness of the category of protocols would mean that there can only be one instance of a note per piece of music. To overcome this, we make our spacetime be $\mathbf{TStamp^{1,14}}$. This automatically allows instructions to refer to sets of notes (and their silences), so we can treat a collection of instructions as one, provided they all have the same duration.

All in all, we can now see music scores as objects in $\mathbf{Prot}(\mathbb{N}, \mathbf{TStamp^{1,14}})$, and a generic event in a protocol here is of the form (n, m_X), where n is an abuse of notation for the identity morphism 1_n in $\mathbb{N}$. The number n says that the note or silence needs

[2]Although the general framework for composition of protocols has not yet been developed, see future work in Chapter 4

to be active for a duration $1/n$ of a whole note. The number m says that this is the mth occurrence of the set of notes and/or silences X, which this instruction refers to. Now only one piece of information is missing: when should the note start to play? For this, we use the happens-before relation, to indicate that a note is to be played right after the previous note or the previous silence is over. This defines a relation whose transitive closure we can take as the happens-before relation. We will assume that, unless otherwise specified, notes are by default in silence. So, instructions for silence can be used when there is a need to specify the start of a sound which is not right after the end of any other sound.

The following short example is designed to demonstrate how this approach allows us to tackle polyphony in which notes start and end at different times, without the need for realistic coordinates.

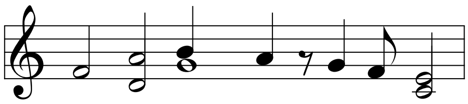


Figure 3.5: a short musical example

Some choices need to be made if we want to see this score as a protocol. For instance, if two notes start and end at the same time, we can regard them as both a single event and as a pair of events. In this example, we group notes together as one event whenever possible. Recall that this is not possible when the notes do not start or end at the same time, since all the notes in an event have only one duration.

With this choice, the protocol that models this music score has been drawn in Figure 3.6. In it, the arrows represent the covering relation for happens-before. Note that the notes on the left side and the dashed lines have been included for visual reference, and that they do not form part of the data of the protocol. Following the translation between objects in the category of protocols and diagrams introduced in 3.5.1, we can equivalently see this music score as the diagram in Figure 3.7. Unlike in the protocol, here the notes and the lines are part of the data for the diagram.

As we will explain in Chapter 4, music is a nice toy example which gives us a simple setting in which to try out and explore different ideas.

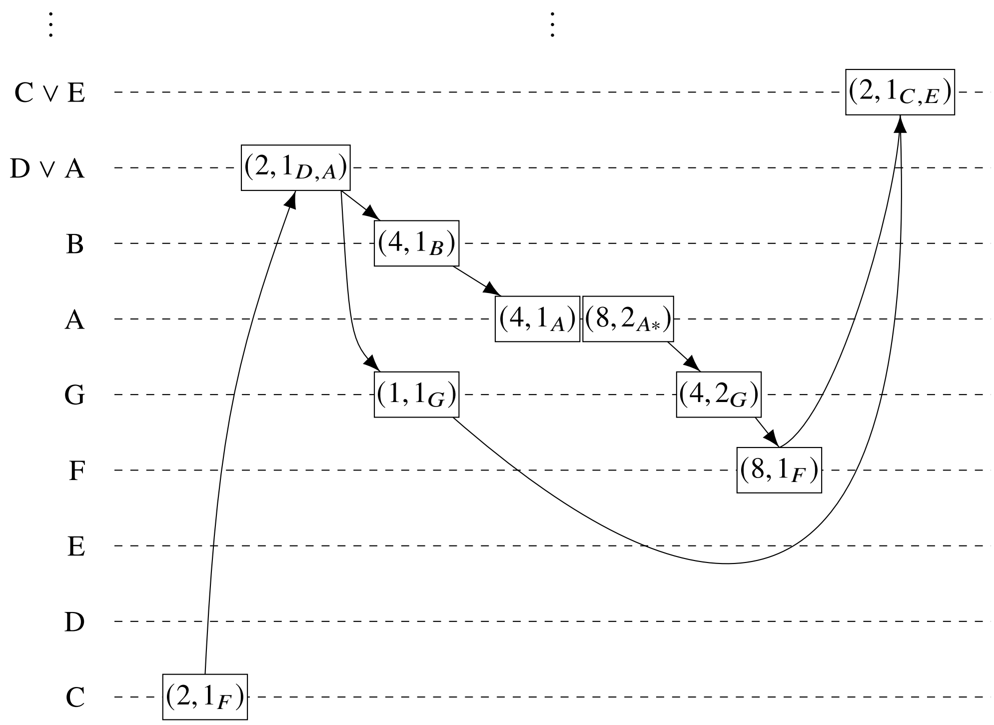


Figure 3.6: Music as a protocol

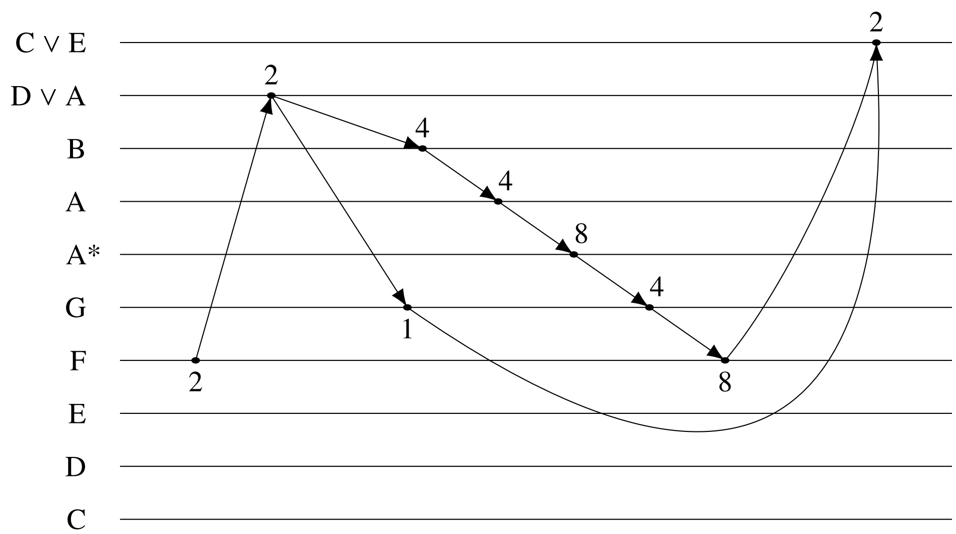


Figure 3.7: music as a diagram

3.5.3 Quantum circuits

The CP construction

To calculate the action of a given quantum circuit, it is commonplace to work only with pure states: each qubit is in a state which can in general be arbitrary, $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, and then after every gate the state gets updated appropriately. Whenever there is a measurement, we talk about the history of that qubit branching. The way in which this branching is felt by other qubits, when they get classically controlled by the outcome of a measurement, cannot be computed as the result of applying one gate to the qubit. After all, measurement is not reversible, so it is not a gate.

Thankfully, however, there is a way to put preparation of qubits, unitary gates and measurement in the same mathematical footing. That means we can regard each of them, and thus their compositions, as morphisms. This is a particular case of the so-called CP construction [43, Chapter 7]. The idea is to switch from working only with pure states to working with mixed states. Next, we give some examples of how to view some of the basic building blocks of quantum circuits as matrices, and thus also as morphisms in $CP[\mathbf{FHilb}]$.

A quantum wire is just the object $M_{2\times2}(\mathbb{C})$, so a quantum state is a two-by-two matrix. We represent a pure quantum state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ via the following density matrix.

$$\rho_\psi = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \beta\alpha^* & |\beta|^2 \end{pmatrix} \tag{3.2}$$

If we only know that the system is in one of the pure states $|\psi_1\rangle, \ldots, |\psi_n\rangle$, with probability λ_i of being in each state $|\psi_i\rangle$, we say that the system is in a mixed state. To represent this as a density matrix, simply compute $\sum_{i=1}^n \lambda_i |\psi_i\rangle\langle\psi_i|$. As we can see, pure states are, from this point of view, put in the same footing as mixed states. The same will happen for gates, measurement, and state preparation.

By the Kraus representation theorem [89, Theorem 2], a linear map F is completely positive if and only if it can be written in the form

$$F(A) = \sum_i B_i A B_i^\dagger \tag{3.3}$$

for some matrices B_i.

As usual, the matrix that represents a controlled-U gate is

$$C(U) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & u_{00} & u_{01} \\ 0 & 0 & u_{10} & u_{11} \end{pmatrix} \tag{3.4}$$

An arbitrary input for a two qubit gate is the Kronecker product of the control qubit and the controlled qubit (in that order), which can also be seen as the density matrix for the arbitrary pure state $|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$, .

When going through a gate U, the state ρ_ψ is transformed into $\rho_{\psi'} = U\rho_\psi U^\dagger$. For instance, when $U = X$ we obtain the density matrix for the state $|\psi'\rangle = X|\psi\rangle$. If $|\psi\rangle$ is a two-qubit state and $U = CNOT$, we obtain the density matrix for $|\psi'\rangle = CNOT|\psi\rangle$.

The operation of measuring a single qubit in the computational basis is the map

$$M(\rho) = |0\rangle\langle 0|\rho|0\rangle\langle 0| + |1\rangle\langle 1|\rho|1\rangle\langle 1| \tag{3.5}$$

which we know is completely positive because of the Kraus representation theorem. This acts by simply setting the non-diagonal elements to 0.

To describe a classically controlled U-gate, we need a matrix $CC(U)$ that takes as input

$$\rho = \begin{pmatrix} |\alpha_1|^2 & 0 \\ 0 & |\beta_1|^2 \end{pmatrix} \otimes \begin{pmatrix} |\alpha_2|^2 & \alpha_2\beta_2^* \\ \beta_2^*\alpha_2 & |\beta_2|^2 \end{pmatrix} \tag{3.6}$$

where $|\psi_i\rangle = \alpha_i|0\rangle + \beta_i|1\rangle$, and such that the output $CC(U)\rho CC(U)^\dagger$ is the density matrix for the mixed state '$|\psi_2\rangle$ with probability $|\alpha_1|^2$ and $U|\psi_2\rangle$ with probability $|\beta_1|^2$'. It is easy to check that this requirement forces, for instance,

$$CC(X) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.7}$$

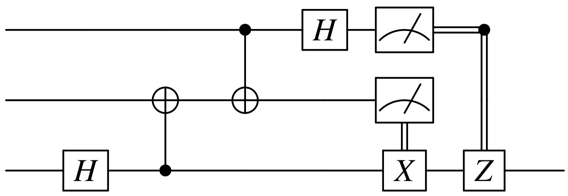
Quantum circuits as protocols

Now we understand that the all the processes (qubit preparations, gates and measurements) that are allowed within quantum circuits are modelled by the morphisms of **CP**[**FHilb**]. If we assign the appropriate timestamps to these possible events, then we can see quantum circuits as protocols, and more concretely as objects in

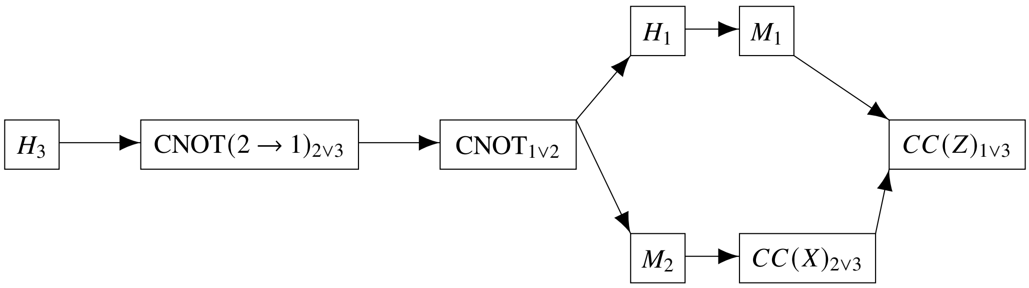
$\mathbf{Prot}(\mathbf{CP}[\mathbf{FHilb}], \mathbf{TStamp^{1,n}})$, where n is the number of qubits involved. Equivalently, we can see circuits diagrammatically as morphisms in $\mathbf{Diag}(\mathbf{CP}[\mathbf{FHilb}])(2^n-1, 2^n-1)$, since 2^n-1 is the number of non-empty subsets of the set of size n, which is the number of possible qubit combinations in which a single gate can happen.

Given a quantum circuit, each gate U will correspond to an event of the form $(M(U), t_i)$ in our protocol, where $M(U)$ is the matrix (morphism in $\mathbf{CP}[\mathbf{FHilb}]$) that represents the action of U, the qubit or collection of qubits on which the gate acts is modelled by i, and t is the Lamport clock for the event, which informs about where to place it within location i. We can denote this event by $U_{(t,i)}$. After translating each gate to an event, we only have to define the happens-before preorder between the events. This is generated by the following covering relation: $U_{(t,i)} \lessdot V_{(r,j)}$ whenever $i \cap j \neq \emptyset$ and V appears to the right of U in the quantum circuit's diagram.

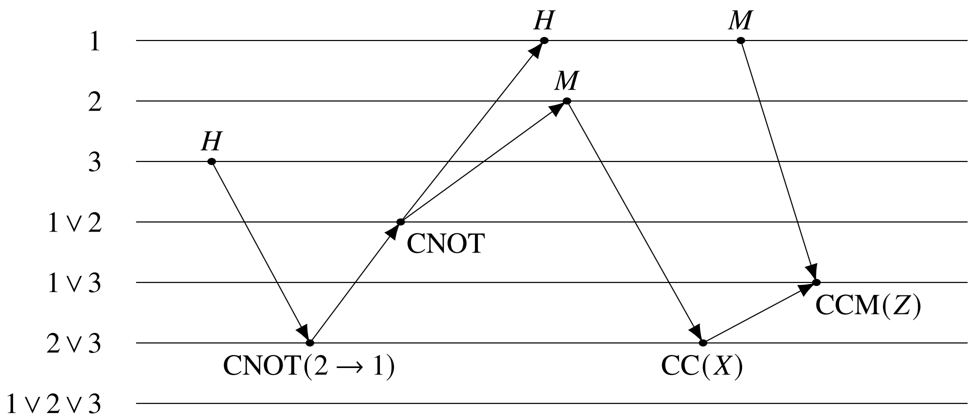
As a concrete example, let us look at the circuit for quantum teleportation.



We can also view this as an object in the category of protocols, as follows.



Finally, we can see it as a the following diagram.



3.6 Interlude: causality

As we explained in the introduction, part of the data of any given spacetime manifold $\mathcal{M}$ is the causal partial order for points, which lets us define the following regions called the future and past cones of $p \in \mathcal{M}$.

$$J^+(p) = \{q \in X \mid p \leq q\}$$
$$J^-(p) = \{q \in X \mid q \leq p\}$$

We may regard $J^+(-)$ and $J^-(-)$ as operators that take a point and return their causal future and past respectively, and we can generalise them to take arbitrary regions as inputs.

Definition 3.21. For $S \subseteq \mathcal{M}$, define its *future cone* as

$$C^+(S) = \bigcup_{s \in S} J^+(s) = \{q \in \mathcal{M} \mid \exists s \in S : s \leq q\}$$

and its *past cone* as

$$C^-(S) = \bigcup_{s \in S} J^-(s) = \{q \in \mathcal{M} \mid \exists s \in S : q \leq s\}$$

In other words, we take the future and past cones of regions to consist of the union of all the cones from their points. We can characterise these operations algebraically as closure operators, and generalise this to get a notion of causal structure in monoidal categories.

Definition 3.22. A *closure operator* on a partially ordered set $(P, \leq)$ is a function $C : P \to P$ such that

$$s \leq t \Rightarrow C(s) \leq C(t)$$

$$s \leq C(s)$$

$$C(C(s)) \leq C(s)$$

A *causal structure* on a firm monoidal category consists of a pair (C^+, C^-) of closure operators on its partially ordered set of subunits. [3]

[3]Note that we make no requirement for any relationship between these two operators in general, because this depends heavily on the specific context or application, that is, on how these operators are to be interpreted and used. It is left for future work to further explore which specific relationships would make sense to require in different contexts.

Proposition 3.23. *[26, Proposition 27] Causal structure restricts: if $r \in \mathrm{ISub}(\mathbf{C})$ and C is a closure operator on $\mathrm{ISub}(\mathbf{C})$ for some firm $\mathbf{C}$, then $D(s) := C(s) \wedge r$ is a closure operator on $\mathrm{ISub}(\mathbf{C}|_r)$.*

Proof. By Proposition 2.30, the idempotent subunits in $\mathbf{C}|_r$ are those subunits $s \in \mathrm{ISub}(\mathbf{C})$ with $s \leq r$. For $s \leq t \leq r$, we have $s \leq C(s)$ so $s \leq D(s)$, and $C(s) \leq C(t)$ so $D(s) \leq D(t)$. Finally, note that $s \leq r \Rightarrow C(s) \wedge r \leq C(s)$, so $C(s) \leq C(C(s)) \leq C(C(s) \wedge r)$, and so $D(D(s)) \leq D(s)$. □

Proposition 3.24. *Both C^+ and C^- in Definition 3.21 are closure operators on the poset of arbitrary regions of a manifold $(\mathcal{P}(\mathcal{M}), \subseteq, \cap)$.*

Proof. Recall that $s \leq s$ for all $s \in X$ because, as defined in the introduction, $\leq$ is a partial order. □

When closure operators are defined on a locale, we can alternatively view them as additional preorders on the locale which satisfy some extra axioms of compatibility with the already present partial order.

Lemma 3.25. *Fix a locale $(P, \leq, \wedge)$. Then, there is a one-to-one correspondence between closure operators C^+ in P and preorders $\sqsubseteq$ in P satisfying*

$$s \leq t, s \sqsubseteq p \Rightarrow t \sqsubseteq p \tag{1}$$

$$\bigvee \{q \mid s \sqsubseteq q\} \sqsubseteq p \Rightarrow s \sqsubseteq p \tag{2}$$

Proof. Given a closure operator C^+ on P, define $s \sqsubseteq t \iff t \leq C^+(s)$. This is clearly a preorder (not antisymmetric in general) satisfying (1). For (2), note that $\bigvee\{q|s \sqsubseteq q\} = C^+(s)$.

Given a preorder $\sqsubseteq$ in P such that (1) and (2) hold, define $C^+(s) = \bigvee\{t|s \sqsubseteq t\}$. First, we want to show that $s \leq t \Rightarrow C^+(s) \leq C^+(t)$. For this, it is enough to show that $\{p \mid s \sqsubseteq p\} \subseteq \{q \mid t \sqsubseteq q\}$, which follows from (1). It is straighforward that $s \leq C^+(s)$. Finally, we want to show that $C^+(C^+(s)) = C^+(s)$. Note that it is enough to show $C^+(C^+(s)) \leq C^+(s)$, since the other inequality follows from $s \leq C^+(s)$. For this, note that axiom (2) implies $\left\{q' \mid \bigvee\{q \mid s \sqsubseteq q\} \sqsubseteq q'\right\} \subseteq \{p \mid s \sqsubseteq p\}$.

To see that this correspondence is bijective, need to check that $s \sqsubseteq_{New} t \Rightarrow s \sqsubseteq t$. Indeed, by (2) it is enough to show that $\bigvee\{q|s \sqsubseteq q\} \sqsubseteq t$, and by (1) it is enough to find $r \in P$ such that $r \leq \bigvee\{q|s \sqsubseteq q\}$ and $r \sqsubseteq t$. But we can choose $r = t$. □

Following the analogy with physics, the preorder is the same as considering the future cone of forward time-like and null curves from a region. We do not need our preorder to exactly mean causal relationship. This would require it to be irreflexive, since events cannot causally influence themselves. Rather, it is enough that we can recover the causal information from the order. In our case, we know that s may causally influence t if and only if $s \sqsubseteq t$ and $s \neq t$.

The bijection in Lemma 3.25 is not unique, and here we present another one that will be useful to us.

Lemma 3.26. *Fix a locale* $(P, \leq, \wedge)$. *Then, there is a one-to-one correspondence between closure operators* C^- *in* P *and preorders* $\sqsubseteq_P$ *in* P *satisfying*

$$s \leq t, q \sqsubseteq_P s \Rightarrow q \sqsubseteq_P t \tag{1'}$$

$$q \sqsubseteq_p \bigvee \{q' \mid q' \sqsubseteq_P s\} \Rightarrow q \sqsubseteq_P s \tag{2'}$$

Proof. Given C^-, define $s \sqsubseteq_P t \iff s \leq C^-(t)$, and given $\sqsubseteq_P$ define the function $C^-(t) = \bigvee \{s \mid s \sqsubseteq_P t\}$. The rest of the proof is analogous to the proof of Lemma 3.25. □

From the closure operators that give the future and past cones of arbitrary regions as in Definition 3.21, we can get a pair of closure preorders between the regions. Since we have two bijections available, we have a choice on how to do this. To make the interpretation more straightforward, we choose the correspondence in Lemma 3.25 for the future cone, and the correspondence in Lemma 3.26 for the past cone. With that in mind, we obtain the following preorders.

Definition 3.27. Define the *future causal preorder* $\preceq_{\text{Future}}$ between spacetime regions $U, V \in \mathcal{P}(\mathcal{M})$ via

$$U \preceq_{\text{Future}} V \iff V \subseteq C^+(U) \iff \forall v \in V \exists u \in U : u \leq v$$

Similarly, the *past causal order* $\preceq_{\text{Past}}$ between regions is defined as

$$U \preceq_{\text{Past}} V \iff U \subseteq C^-(V) \iff \forall u \in U \exists v \in V : u \leq v$$

In other words, we say that a region V is in the causal future of U if every point in V can be causally influenced by some point in U, and that U is in the causal past of V if every point in U causally precedes some point in V.

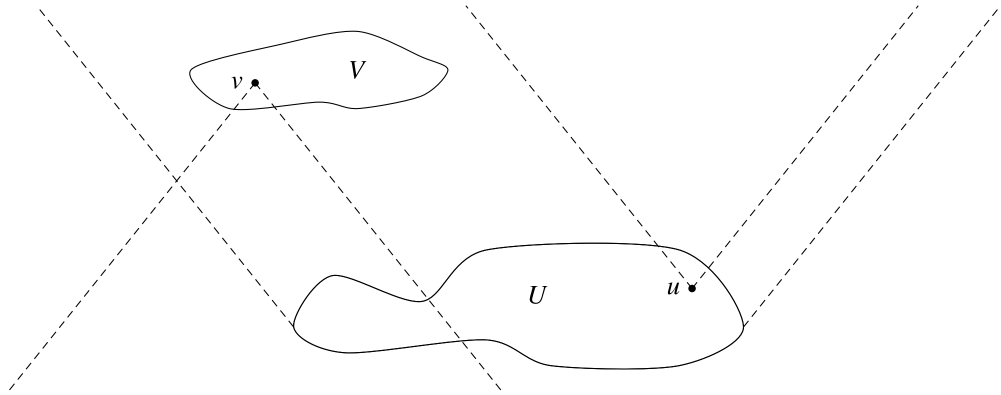


Figure 3.8: Causal order for regions. Note how $U \preceq_{\text{Future}} V$ but $U \not\preceq_{\text{Past}} V$. All the dashed lines represent light rays, and as in Figure 1.1, time is in the vertical axis

This next lemma shows that the axioms (1), (2), (1') and (2') are not unreasonable, since they are automatically satisfied for the free locale on a semilattice.

Lemma 3.28. *Consider a semilattice X together with an additional preorder $\sqsubseteq$ on X. Recall that the free locale is then $(\mathrm{Down}(X), \subseteq, \cap)$, where $\mathrm{Down}(X) = \{S \subseteq X \mid \downarrow S = S\}$. Define the following preorders based on the given one $\sqsubseteq$.*

$$U \sqsubseteq_F V \iff \forall v \in V \exists u \in U : \quad u \sqsubseteq v$$
$$U \sqsubseteq_P V \iff \forall u \in U \exists v \in V : \quad u \sqsubseteq v$$

Then, $\sqsubseteq_F$ is a preorder in $\mathrm{Down}(X)$ satisfying axioms (1) and (2), and $\sqsubseteq_P$ is a preorder satisfying axioms (1') and (2').

Proof. This follows straightforwardly from the definitions, without need to use the fact that the elements of the locale are downwards closed. □

Let us finish this section by comparing causal structure for regions to the usual one for points. In the case of points in spacetime, we can say

$$x \le y \iff y \in J^+(x) \iff x \in J^-(y) \tag{3.8}$$

So, future and past operators can be induced from each other. The next lemma shows that, whenever the causal preorder satisfies axioms (1), (2), (1') and (2'), this remains possible in general.

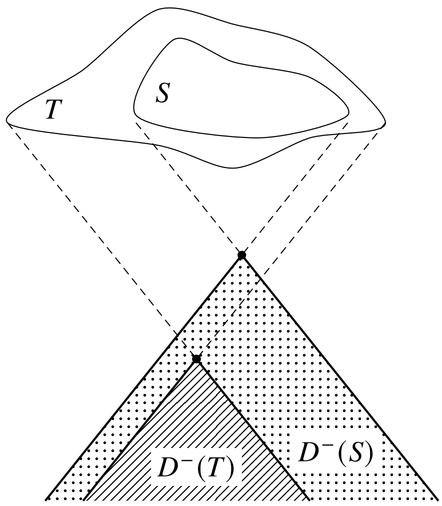
Lemma 3.29. *Fix locale $(P, \le, \wedge)$ and a preorder $\sqsubseteq_F$ satisfying axioms (1) and (2), and consider closure operator $C^+(x) = \bigvee\{q \mid x \sqsubseteq_F q\}$ in P. Then, $x \sqsubseteq_F y \iff y \le C^+(x)$. Similarly, if we consider a preorder $\sqsubseteq_P$ satisfying axioms (1') and (2'), and closure operator $C^-(y) = \bigvee\{q \mid q \sqsubseteq_P y\}$, then $x \sqsubseteq_P y \iff x \le C^-(y)$.*

Proof. Since the proof of these two facts is similar, we will focus on the first one. If $x \sqsubseteq_F y$, then $y \in \{q \mid x \sqsubseteq_F q\}$ so $y \leq C^+(x)$. Conversely, assume that $y \leq \bigvee\{q \mid x \sqsubseteq_F q\}$. To show $x \sqsubseteq_F y$, by (2) it is enough to show $\bigvee\{q \mid x \sqsubseteq_F q\} \sqsubseteq_F y$, and by (1) it is enough to find $r \in P$ such that $r \leq \bigvee\{q \mid x \sqsubseteq_F q\}$ and $r \sqsubseteq_F y$. But we can simply choose $r = y$. □

However, and unlike in the usual causal structure for points, a causal preorder for regions which satisfies axioms (1) and (2) and can induce a future closure operator is not, in general, able to induce a closure operator for the past, and vice-versa.

Lemma 3.30. *Fix a locale $(P, \leq, \wedge)$ and a preorder $\sqsubseteq_F$ in P satisfying axioms (1) and (2). Then, in general the function defined on P as $D^-(x) = \bigvee\{q \mid q \sqsubseteq_F x\}$ is not a closure operator. Similarly, if we have a preorder $\sqsubseteq_P$ satisfying axioms (1') and (2'), the function $D^+(x) = \bigvee\{q \mid x \sqsubseteq_P q\}$ is not, in general, a closure operator.*

Proof. Take $\mathcal{M} = \mathbb{R}^2$, and draw light rays as dashed lines as in Figure 3.8 above. If we take $\preceq_{Future}$ as in Definition 3.21, and define $D^-(U) = \bigvee\{Q \mid Q \preceq_{Future} U\}$, we want to show that this is not a closure operator. In particular, consider the following situation.



Note that $D^-(S)$ is the union of all regions that could causally influence every point in S, so in the picture this corresponds to the dotted area. Similarly, $D^-(T)$ corresponds to the dashed area, and we can see that, even though $S \subsetneq T$, we have $D^-(T) \subsetneq D^-(S)$. So, D^- is not a closure operator. The proof for the second statement can be obtained by turning this figure upside down. □

Finally, we can now describe whether or not a causal dependency between events, expressed by the happens-before relation, is realisable in spacetime.

Definition 3.31. Suppose that $\mathbf{D}$ has a causal structure. Then, we say that a protocol $A \in \mathbf{Prot}(\mathbf{C}, \mathbf{D})$ is *compatible* with the causal structure of $\mathbf{D}$ if

$$(f, s) \sqsubseteq_A (g, t) \Rightarrow t \leq C^+(s)$$

Intuitively, compatibility requires every process to take place within the causal future of all events that happen before it, as expected according to our interpretation of the happens-before preorder as events having to wait for others. Since the protocols compatible with the causal structure of $\mathbf{D}$ are closed under tensoring, we can see them as forming a monoidal subcategory.

3.7 Examples with realistic model of space and time

In this section, we continue our programme to add more details to our categorical models of protocols. We have already seen how labels constitute a toy model for spacetime, formalised by the category of timestamps $\mathbf{TStamp^{n,m}}$, and some examples of protocols as objects in $\mathbf{Prot}(\mathbf{C}, \mathbf{TStamp^{n,m}})$ for some $\mathbf{C}$, n and m. Now, we want to add more details by having more realistic points or regions of spacetime as timestamps. As we explained in Section 3.4, these can be formalised as $\mathcal{P}(\mathcal{M}) = (\mathcal{P}(\mathcal{M}), \subseteq, \cap)$ for a given spacetime manifold $\mathcal{M}$. So, the goal of this section is to look at some examples of protocols as objects in $\mathbf{Prot}(\mathbf{C}, \mathcal{P}(\mathcal{M}))$ for some $\mathbf{C}$. Note that, since we only aim to illustrate how the category of protocols can be used, it will be enough to leave the manifold as generic.

Recall that an object $A \in \mathbf{Prot}(\mathbf{C}, \mathbf{D})$ consists of a finite collection of events (f, s) and a happens-before preorder $\sqsubseteq_A$. Generally, we use the preorder to inform about the causal structure of the protocol, or in other words about which events have to wait for which others, and an event (f, s) specifies that process f is to happen within region s.

3.7.1 Teleportation

Suppose that Alice holds an unknown quantum state ϕ, which she wants to send to Bob. If this was classical information, she could measure ϕ and create a copy. Upon measuring quantum information, however, the state gets destroyed and cannot be reproduced. Also, the no-cloning theorem makes it impossible to systematically copy unknown quantum states. Despite this, it is still possible to send unknown quantum states by using entanglement as a resource, by using the teleportation protocol.

In it, the first step is to create an entangled pair and to send half of it to Alice and the other half to Bob. Secondly, Alice measures her half together with ϕ in a Bell measurement, obtaining one out of four possible results. This result needs to be classically sent to Bob, for it will allow him to choose which correction to perform on his half, and thus to obtain the state ϕ as wanted.

Abstractly, teleportation can occur in an arbitrary monoidal category **C** with right duals [43, Section 3.2], and we will focus on this since the usual teleportation of quantum states is the particular case **C** = **FHilb**. As usual in the CQM modeling of protocols, each step involved is represented by a morphism, and by composing and tensoring these morphisms appropriately we get a new morphism that models the whole protocol.

In our setting, pair creation will be an event (η, c). Alice holding state ϕ is the event $(1_\phi, a)$, and performing the measurement is $(\varepsilon \circ (1 \otimes U^*), a')$, where U^* is the dual of the correction that Bob will need to apply. Finally, Bob performing the correction is the event (U, b). As for the happens-before preorder, see the Hasse diagram in Figure 3.10.

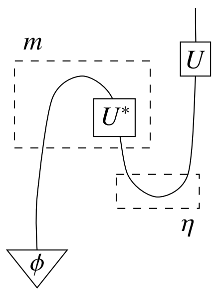


Figure 3.9: teleportation in CQM

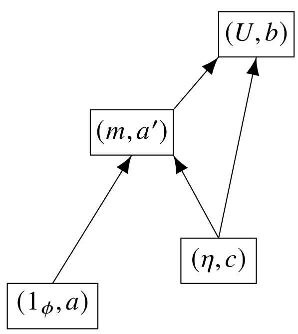


Figure 3.10: teleportation as an object in the category of protocols

Note how m needs to happen-before Bob's correction, unlike in the CQM version where these are tensored so they can graphically appear either above or below each other. Furthermore, by asking this protocol to be compatible with the causal structure of spacetime, we can rule out situations such as having the region for the correction in the causal past of the region for pair creation.

If we are willing to lose detail in the regions of spacetime, we can replace the regions $a, b, c, a' \in \mathcal{P}(\mathcal{M})$ by some labels. For instance, a and a' may represent Alice's lab at different points in time, so we could abstract this as one of the parties involved. Similarly, we label c as Charlie's lab and b as Bob's lab. The advantage of having timestamps be labels, as we discussed in Section 3.5, is that we can label the messages as such. Being able to distinguish messages from events allows us to interpret them differently, as the following example shows.

Example 3.32. In the following figure, messages are allowed to carry their own timestamps. For instance, the messages that are sent from the pair creating event contain both sides of the pair, and the message from the measurement carries the correction. The correction event in the end means that U is acting on and transforming the qubit, whereas the message from the measurement simply means that U is its content that the receiver will get.

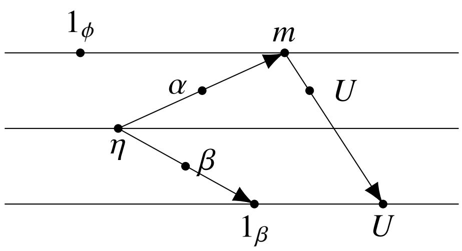


Figure 3.11: teleportation protocol as a diagram modelling the contents of messages

3.7.2 Summoning

There are some protocols that are both relativistic and quantum, in the sense that their very definition requires talking about regions of spacetime as well as about quantum information. Because categorical semantics for protocols was lacking natural ways to refer to locations until now, it was also lacking the ability to model this kind of protocols. Indeed, modelling protocols categorically always results in the locations of events being forgotten. We can now tackle this problem, and we will look at the task of summoning information in spacetime [39] to illustrate how.

Two parties, Alice and Bob, are involved in this task. At some point s, Bob shares with Alice a quantum state ϕ that is known to him but not to her. A finite number of agents A_i are distributed in spacetime, and one of them will request that the state be revealed to him. For each agent, there is a request point y_i and a reveal point z_i. We say that Alice can complete the task successfully if, without knowing in advance which agent will be requesting, she can present the state ϕ at the appropriate reveal point. For simplicity, we will focus on the case with two agents, since once this is solved all other cases can be solved inductively [39].

There are two limitations to Alice's capabilities: information cannot travel faster than the speed of light, and unknown quantum states cannot be systematically copied. However, quantum information can still be 'copied in time', for instance thanks to the teleportation protocol, and thus in spacetime, to an extent that is exactly characterised by the following theorem.

Theorem 3.33. *(No-summoning) [39, Theorem 1] Summoning is possible if and only if*

- *Every reveal point is in the future light cone of the starting point s.*
- *Every pair of causal diamonds is causally related.*

In our setting, the first condition is modelled by requiring that the protocol must be compatible with the causal structure of the spacetime. That means that we need to ask all of the (potentially) revealing events to happen after both the starting event and the corresponding request points. This is enough to model the limitation of no superluminal communication.

The second condition in the theorem ensures that we are not breaking the second limitation, since, as argued in [39], otherwise there would be a strategy to violate the no-cloning theorem. To phrase this a bit more formally in our setting, let us start by recalling that, in categorical semantics, a general version of the no-cloning theorem exists [43, Section 4.2].

Definition 3.34. A braided monoidal category has *uniform copying* if there is a natural transformation $d_A : A \to A \otimes A$ such that $d_I = \rho_I^{-1}$ satisfying the equations in Figures 3.12, 3.13 and 3.15 for all objects A, B. Naturality and $d_I = \rho_I^{-1}$ are graphically represented in Figures 3.16 and 3.14 respectively.

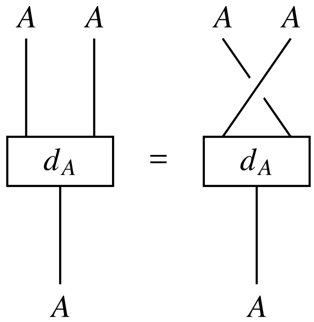


Figure 3.12

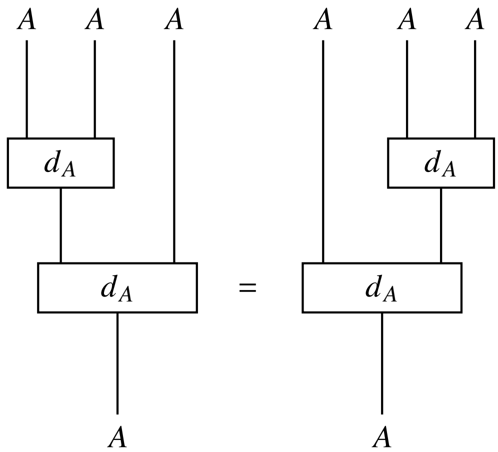


Figure 3.13

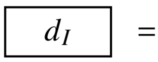


Figure 3.14

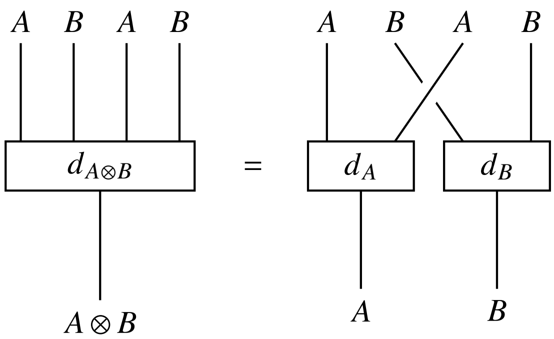


Figure 3.15

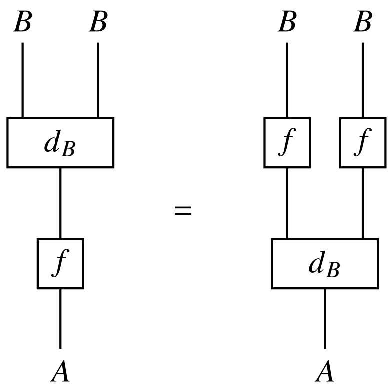


Figure 3.16

Theorem 3.35. *(No-cloning) [43, Theorem 4.17] If a braided monoidal category with duals has uniform copying, then every endomorphism is a scalar multiple of the identity.*

This means that a category is somewhat degenerate (although not necessarily trivial), whenever it has universal copying. We can see the usual no-cloning theorem of quantum mechanics as a particular case of this: because we know that the category **Hilb** has endomorphisms other than multiples of the identity, we know that it cannot have universal cloning.

In our framework, because of how we defined the category of protocols, this always has universal cloning, so we have to use some other strategy if we want to rule out configurations for summoning that break the second condition in Theorem 3.33.

We can now reduce the proof that the second condition in the no-summoning theorem is necessary to the categorical no-cloning theorem. The idea is that, in a configuration in which two diamonds are not causally related, we are realising a morphism of type $A \to A \otimes A$ in $\mathbf{C}$, since that setup allows in principle to teleport the same state twice in parallel. Further, the same configuration with input B would be realising a morphism of type $B \to B \otimes B$ in $\mathbf{C}$, and so on.

Remark 3.36. Following the argument above, we will assume that a configuration for summoning with two causally unrelated diamonds acts as a witness that the category $\mathbf{C}$ has universal cloning.

It follows that such configurations are only allowed when $\mathbf{C}$ is degenerate, so in particular when $\mathbf{C} = \mathbf{FHilb}$ then summoning cannot work in such a configuration. In a way, we can see this as the result of lifting the no-cloning property from $\mathbf{C}$ to $\mathbf{Prot}(\mathbf{C}, \mathbf{D})$.

Chapter 4

Future work

Other constructions, alternative to the category of protocols. We have introduced the category of protocols as one construction that finds applications for the theory of subunits. But there could be other ways to do this, as we have made a number of choices along the way. In particular, we have defined protocols as the objects in the category, because it was not clear how to see them as morphisms. The main advantage of this is that we do not need to worry about what should the domain and codomain of a protocol be. Another advantage is that we get more freedom to choose the morphisms of the category, which in turn lets us define a monoidal structure that yields interesting notions of subunits, restriction and support. However, there is a clear disadvantage: the connection with CQM and even with the category of diagrams $\mathbf{Diag}(\mathbf{C})$ cannot be seen as a functor, since in those settings protocols are morphisms rather than objects. Having protocols as morphisms would also force us to formalise how to compose them.

There is one more source of motivation to look for alternative constructions. At the moment, two protocols are isomorphic if and only if their networks are equal. This means that, as long as the events take place in the same regions, two protocols will be isomorphic even if their causal structures are different. Perhaps we could have a different construction with a notion of isomorphism that only identifies protocols with the same causal structure. More generally, there could be other constructions where the morphisms have more meaning; for instance they could represent translations between protocols which are the same but expressed for different frames of reference.

In CQM, the tensor product means, for objects, that we are considering the composite system, and for morphisms, that we are looking at the parallel composition of processes. In categories whose objects are sets of events, such as the category of protocols, if the tensor product is some kind of union then the tensor unit will be trivial, and there will be no interesting subunits. Therefore, we choose a tensor unit which is more like an intersection. It may be possible to construct a category of this kind with a clear interpretation of the tensor product as parallel composition, and at the same time interesting subunits, but it is not clear how.

Graphical calculus. The graphical notation for monoidal categories that is surveyed in [81] and used extensively in CQM [43, 21] is a very useful tool, since it allows us to reason and compute in much more intuitive ways than the usual mathematical notation based on writing symbols one after the other in a line. Since we have introduced some categorical notions such as restriction, support, and causal structure, we wonder how they can be incorporated into this graphical notation. This may seem a rather unimportant goal, but a successful integration with this notation would mean a better chance for the notions introduced in this thesis to become standard in CQM.

Capturing other properties of protocols. We have seen how the category of protocols can easily detect both the possibility of deadlock as well as the occurrence of deadlock. This is possible because enough information about the protocol is being captured in our model, and for the same reason it should be possible to characterise other properties of distributed protocols, such as livelock or race conditions, as well. In other words, what is the modelling potential of the category of protocols?

Subunits as logical statements. We usually think of subunits as regions of spacetime, since this was our initial motivation. However, as we discussed in Section 2.2, subunits can have other meanings. Recall that we can always, given a semilattice S, find a category whose subunits are exactly S: we can view S as a monoidal category, and also consider its product with any other simple category. What would it mean to have protocols where the subunits have meanings other than regions? For instance, we could consider the well-formed formulas of a formal languge as a semilattice (with *and* as meet and *implies* as order), and view this as a category. Then, an event (f, s) could be seen as labeling a process f with a statement s.

This would demonstrate that, although the theory of subunits was inspired by thinking of regions of spacetime, it is a useful way to encode more information about protocols which goes beyond regions and causal order.

Connection with topological semantics. Modal logic [10] is an extension of classical propositional logic to include a unary modal operator $\Diamond$ (read 'diamond'). When applied to a well-formed formula ϕ, the result $\Diamond\phi$ can be understood as 'it is possible that ϕ'. The concept of necessity can be expressed via the derived unary operator $\Box\phi := \neg\Diamond\neg\phi$, which is then read as 'necessarily ϕ'.

Topological semantics give a model of modal logic via a valuation function that maps formulas to the region where they hold [88, 5]. If a formula ϕ corresponds to some region under the valuation function, then $\Box\phi$ and $\Diamond\phi$ are valuated in the interior and the closure of that region, respectively.

Can this be adapted to our setting, by replacing the codomain of the valuation function with the set of subunits? This could result in categorical semantics for modal logic.

Connection with causality via terminality. Another natural open problem is to establish the relationship between our way to model causality and the other existing ways in the literature, and in particular the one based in terminality as introduced in Section 1.3. We can speculate that it may be possible to recover the causal characterisation of processes based on terminality as a corollary of ours, since broadly speaking we have access to more detailed causal information. In other words, because we are able to discuss causality in greater detail, characterising in particular which processes respect causal structure or not should be possible within our framework, and this characterisation may correspond to the one based on terminality. One possible strategy would be to exploit further the idea of having protocols be compatible with the causal structure of spacetime (that is, having happens-before imply causal precedence, see Definition 3.31).

Connection with event structures. A motivation for trying to relate our formalism with the theory of event structures would be that this could in turn result in a connection with other models of concurrent computation such as Petri nets and Scott domains, very much in the same way that event structures are related to these.

One way to approach this could be to formalise the idea of finding the 'underlying' event structure of an arbitrary object in $\mathbf{Prot}(\mathbf{C}, \mathbf{D})$. Since we already have, in any given protocol, a set of events and their causal structure, given by the happens-before relation, we would only need to define the incompatibility relation. If there is no natural way to induce an incompatibility relation from the data of an arbitrary protocol, it may be possible to simply add this relation as an extra piece of data in the definition of protocol.

Recovering existing models of distributed computation. The category of protocols $\mathbf{Prot}(\mathbf{C},\mathbf{D})$ gives us a scheme for constructing categories, one for each choice of $\mathbf{C}$ and $\mathbf{D}$, and we can see each of these as a particular model for distributed computation. With the appropriate choices of $\mathbf{C}$ and $\mathbf{D}$, is it possible to recover any of the existing models for distributed computation?

Phase space formalism. In physics, a very common approach when modelling a system is to represent all its possible states as points of a *phase space*, which is typically a manifold. The idea is that the space will have one dimension per degree of freedom of the system, and a curve in the space represents a dynamical transformation of the system's state (since we think of the parameter of the curve as time).

So far, with subunits we have captured intuitions from topology, and in particular from spacetimes. Perhaps some ideas from the phase space formalism can be captured in the theory of subunits as well. If so, we wonder to what degree the phase space point of view on physics can coexist or is compatible with the categorical one, where systems are objects and the transformations between them are morphisms.

Applications of the locale-based completion theorem. At the moment, Theorem 2.80 is an interesting theoretical result but more work is required to see how to take advantage of it in applications. Broadly speaking, one motivation to use this result would be as follows. If the subunits of a category only form a semilattice, every morphism is in principle supported on a multitude of subunits. However, when subunits form a frame, we can talk about *the support* of a morphism, that is, the smallest subunit where it is supported.

Technical directions. Can $\mathbf{Hilb}_{\mathbf{C_0(X)}}$ be regarded as the category of Hilbert spaces internal to some ambient category, whose categorical logic governs supports? In general, what other properties of a space X can be seen categorically?

It would be interesting to examine what happens to subunits under constructions such as Kleisli categories, Chu spaces, or the Int-construction [51]. One could ask how much of the theory carries over to skew monoidal categories [84], and how these notions relate to restriction categories [35]. Finally, it would be desirable to find global conditions on a category providing its subunits with further properties, such as being a compact frame or Boolean algebra, or with further structure, such as being a metric space.

Appendix A

Day convolution

This appendix describes in some detail the monoidal structure on presheaf categories given by Day convolution [24], so that it can prove some of the lemmas of Section 2.9. We start with the abstract definition, then give a concrete description, and use that to write down the coherence isomorphisms; we have no need for associators or the braiding in this article, so will not discuss these explicitly. Fix a monoidal category $\mathbf{C}$, and write $\widehat{\mathbf{C}} = [\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$ for the category of presheaves.

Tensor product of objects

The Day convolution $F \mathbin{\widehat{\otimes}} G$ of presheaves $F, G \in \widehat{\mathbf{C}}$ is given abstractly as a left Kan extension

$$F \mathbin{\widehat{\otimes}} G \simeq \mathrm{Lan}_{\otimes}(F \times G)$$

of the functor $F \times G \colon (\mathbf{C} \times \mathbf{C})^{\mathrm{op}} \to \mathbf{Set}$, given by $(A, B) \mapsto F(A) \times G(B)$ and $(f, g) \mapsto F(f) \times G(g)$, along the tensor product $\otimes \colon (\mathbf{C} \times \mathbf{C})^{\mathrm{op}} \to \mathbf{C}^{\mathrm{op}}$ of the base category. This left Kan extension may be computed [66, X.4.1] as a coend

$$(F \mathbin{\widehat{\otimes}} G)(A) = \int^{B,C} \mathbf{C}(A, B \otimes C) \times F(B) \times G(C).$$

Now, coends can be computed as colimits [66, IX.5.1], and in turn, colimits can be constructed from coproducts and coequalizers [66, V.2.2]. Thus $F \mathbin{\widehat{\otimes}} G$ is a coequalizer of the following two functions.

$$\coprod_{\substack{f \colon B \to B' \\ g \colon C \to C'}} \mathbf{C}(A, B \otimes C) \times F(B') \times G(C') \rightrightarrows \coprod_{B,C} \mathbf{C}(A, B \otimes C) \times F(B) \times G(C)$$

$$(h, x, y)_{(f,g)} \mapsto \big((f \otimes g) \circ h, x, y\big)_{(B',C')}$$

$$(h, x, y)_{(f,g)} \mapsto \big(h, F(f), G(g)\big)_{(B,C)}$$

Finally, coproducts in **Set** are disjoint unions, and coequalizers are quotients. Thus

$$(F\widehat{\otimes}G)(A)=\Big(\coprod_{B,C}\mathbf{C}(A,B\otimes C)\times F(B)\times G(C)\Big)/\sim,$$

where $\sim$ is the least equivalence relation satisfying

$$(h,x,y)_{(B,C)}\sim(h',x',y')_{(B',C')}$$

when there exist $f\colon B\to B'$ and $g\colon C\to C'$ such that $x=F(f)(x')$, $y=G(g)(y')$ and $(f\otimes g)\circ h=h'$.

$$\begin{tikzcd} & A \arrow[dl, "h"'] \arrow[dr, "h'"] & \\ B\otimes C \arrow[rr, "f\otimes g"'] & & B'\otimes C' \end{tikzcd}$$

It also follows that the action of $F\widehat{\otimes}G$ on a morphism $f\colon A'\to A$ is given by $(h,x,y)_{(B,C)}\mapsto(h\circ f,x,y)_{(B,C)}$.

Tensor product of morphisms

If $\varphi\colon F\Rightarrow F'$ and $\psi\colon G\Rightarrow G'$ are natural transformations, then so is $\varphi\widehat{\otimes}\psi\colon F\widehat{\otimes}G\Rightarrow F'\widehat{\otimes}G'$, given by

$$(\varphi\widehat{\otimes}\psi)_A\colon(h,x,y)_{(B,C)}\mapsto(h,\varphi_B(x),\psi_C(y))_{(B,C)}.$$

Tensor unit

If I is the tensor unit of $\mathbf{C}$, then $\widehat{I}=\mathbf{C}(-,I)$ is the tensor unit of $\widehat{\mathbf{C}}$.

Unitors

Write $\rho_A\colon A\otimes I\to A$ and $\lambda_A\colon I\otimes A\to A$ for the unitors in $\mathbf{C}$. The right unitor $\widehat{\rho}_F\colon F\widehat{\otimes}\widehat{I}\Rightarrow F$ is given by

$$(\widehat{\rho}_F)_A\colon(h,x,y)_{(B,C)}\mapsto F\big(\rho_B\circ(1_B\otimes y)\circ h\big)(x).$$

and the left unitor $\widehat{\lambda}_F\colon\widehat{I}\widehat{\otimes}F\Rightarrow F$ by

$$(\widehat{\lambda}_F)_A\colon(h,x,y)_{(B,C)}\mapsto F\big(\lambda_C\circ(x\otimes 1_C)\circ h\big)(y).$$

It is straightforward to check that these are well-defined natural isomorphisms.

Subunits

A subunit S is firstly a subobject of $\widehat{I}$, i.e. a subfunctor of $\mathbf{C}(-,I)$. Equivalently, to each object A it assigns a set $S(A)$ of morphisms $A \to I$, and naturality amounts to these being closed under precomposition with arbitrary morphisms of $\mathbf{C}$, i.e. whenever $s \in S(A)$ and $f\colon B \to A$ then $s \circ f \in S(B)$. Finally S being a subunit means precisely that for all $s \in S(A)$ there exists a unique $(h,x,y)_{(B,C)} \in (S \widehat{\otimes} S)(A)$, for some $h\colon A \to B \otimes C$, $x \in S(B)$, and $y \in S(C)$, with $s = \rho_I \circ (x \otimes y) \circ h$.

Proof of Example 2.72

By the above description, subunits in $\widehat{M}$ correspond to ideals $S \subseteq M$ which are idempotent in the sense that $S = SS$, and furthermore satisfy the requirement that the map $S \widehat{\otimes} S \to S$ is injective.

Let S be the ideal consisting of all elements of the form $(a,0)+x$ for some $a > 0$, and T the ideal of all elements of the form $(0,b)+y$ for $b > 0$, similarly. We claim that these are subunits. If $\widehat{M}$ were firm, then $S \widehat{\otimes} T = S \cap T$ being a subunit and hence idempotent as an ideal. But $S \cap T$ is not idempotent.

Indeed, consider $(0,1) \in S \cap T$. Now suppose that $(0,1) = (a,b)+(c,d)$ for some $(a,b),(c,d) \in S \cap T$. Then $a+b+c+d = 1$. If $a+c < 1$ necessarily $a = c = 0$. Now $b > 0$ or $d > 0$, so either $b < 1$ or $d < 1$; without loss of generality say $b < 1$. But this contradicts $(a,b) \in S$. Therefore $a+c = 1$. But then $b = d = 0$, contradicting $(a,b) \in T$. Thus $S \cap T$ is not idempotent.

It remains to verify that S and T are subunits. We first treat the case for S. Firstly, S is idempotent since each element $(a,0)$ for $a > 0$ has $(a,0) = (a/2,0)+(a/2,0)$ with $(a/2,0) \in S$. Finally, we must check that any $(h,s,t) \in S \widehat{\otimes} S$ is determined by its value $hst \in M$.

Note that $(h,s,t) \sim (h+x+y, s', t')$ when $s = s'+x$ and $t = t'+y$ for $h,x,y \in M$ and $s,s',t,t' \in S$. Hence any element (h,s,t) is equivalent to one of the form $\big((b,c),(a,0),(a,0)\big)$ for arbitrarily small $a > 0$. Now suppose that

$$(b,c)+(a,0)+(a,0) = (b',c')+(a',0)+(a',0)$$

Using the same trick again we may assume that $a' = a$. Now if $b+a+a > 1$ there is some $d < a$, say with $a = d+e$, such that $b+d+d > 1$ also. Letting $(b',c') = (b,c)+(2d,0)$ gives $\big((b,c),(a,0),(a,0)\big) = \big((b',c'),(e,0),(e,0)\big)$, now with $b'+e+e < 1$. Applying this trick we may have assumed to begin with that $b+a+a < 1$ and $b'+a+a < 1$. But this ensures that $b = b'$ and $c = c'$, and we are done. Seeing that T is a subunit is similar but simpler. $\square$

Proof of Lemma 2.75

As noted when proving Proposition 2.76, we may assume that U and V are idempotent. The lower isomorphism of (2.8) follows from monoidality of the Yoneda embedding.

By definition, $\big(\langle U,X\rangle \widehat{\otimes} \langle V,Y\rangle\big)(A)$ consists of triples (h,f,g) where $h\colon A \to B\otimes C$, $f\colon B \to X$ restricts to U, and $g\colon C \to Y$ restricts to V, subject to the Day identification rules. From the definition of the monoidal structure in $\widehat{\mathbf{C}}$, the transformation $u \widehat{\otimes} v$ in (2.8) has component at A given by

$$(h,f,g) \mapsto ((f\otimes g)\circ h\colon A \to X\otimes Y)$$

Since this is well-defined and each such morphism $(f\otimes g)\circ h$ clearly restricts to a member of $U\otimes V$, it restricts to a transformation as in the top row of (2.8), making the diagram commute. Furthermore each such map is surjective since any morphism $k\colon A \to X\otimes Y$ restricting to a member of $U\otimes V$ has, using the braiding, that $k = ((s\otimes X)\otimes(t\otimes Y))\circ h$ for some h, $s\in U$ and $t \in V$ so that $(h, u\otimes X, v\otimes Y) \mapsto k$.

Finally, we show injectivity. For any triple (h,f,g) with $f = (s\otimes X)\circ \bar{f}$ and $g = (t\otimes Y)\circ \bar{g}$ for some $\bar{f},\bar{g}, s\in U$ and $t\in V$,

$$(h,f,g) \sim ((\bar{f}\otimes\bar{g})\circ h, s\otimes X, t\otimes Y)$$

by the Day identification rules, and so it suffices to consider triples of this form. Now if $(h, s\otimes X, t\otimes Y)$ and $(h', s'\otimes X, t'\otimes Y)$ are mapped to the same morphism then it restricts to $s\wedge s' \in U$ and $t\wedge t' \in V$, so that for some k:

$$\begin{aligned}(k,(s\wedge s')\otimes X,(t\wedge t')\otimes Y) &\sim (h, s\otimes X, t\otimes Y)\\ &\sim (h', s'\otimes X, t'\otimes Y)\end{aligned}$$

by definition of $\sim$, making these triples equivalent as required. $\square$

Bibliography

[1] M. Steen A. Tanenbaum. *Distributed systems : principles and paradigms*. Pearson Prentice Hall, second edition.. edition, 2007.

[2] S. Abramsky and B. Coecke. A categorical semantics of quantum protocols. In *Symposium on Logic in Computer Science*, volume 19, pages 415–425, 2004.

[3] J. Adámek and J. Rosicky. *Locally presentable and accessible categories*. Cambridge University Press, 1994.

[4] S. Ahmad. Group theory, the new language of modern physics. Technical report available at `https://inis.iaea.org/collection/NCLCollectionStore/_Public/22/052/22052503.pdf`, 1990.

[5] S. Awodey and K. Kishida. Topology and modality: The topological interpretation of first-order modal logic. *The Review of Symbolic Logic*, 1(2):146–166, 2008.

[6] P. Balmer. The spectrum of prime ideals in tensor triangulated categories. *Journal für die Reine und Angewandte Mathematik*, 588:149–168, 2005.

[7] P. Balmer and G. Favi. Generalized tensor idempotents and the telescope conjecture. *Proceedings of the London Mathematical Society*, 102(6):1161–1185, 2011.

[8] P. Balmer, H. Krause, and G. Stevenson. The frame of smashing tensor-ideals. *Mathematical Proceedings of the Cambridge Philosophical Society*, 2018.

[9] B. Banaschewski. Another look at the localic Tychonoff theorem. *Commentationes Mathematicae Universitatis Carolinae*, 29(4):647–656, 1988.

[10] P. Blackburn. *Modal logic*. Cambridge tracts in theoretical computer science ; 53. Cambridge University Press, 2001.

[11] R. Blute and M. Comeau. Von neumann categories. *Applied Categorical Structures*, 23(5):725–740, 2015.

[12] F. Borceux. *Handbook of Categorical Algebra 1: Basic Category Theory*. Cambridge University Press, 1994.

[13] F. Borceux. *Handbook of Categorical Algebra 3: Categories of Sheaves*. Cambridge University Press, 1994.

[14] F. Borceux and C. Quinteiro. A theory of enriched sheaves. *Cahiers de Topologie et Geometrie Differentielle Categoriques*, XXXVII-2:145–162, 1996.

[15] M. Boyarchenko and V. Drinfeld. Idempotents in monoidal categories. Available at `bit.ly/2OJichB`, 2009.

[16] M. Boyarchenko and V. Drinfeld. Character sheaves on unipotent groups in positive characteristic: foundations. *Selecta Mathematica*, 20(1):125–235, 2014.

[17] T. Bradley. What is applied category theory? *arXiv:1809.05923*, 2018.

[18] C. Cassidy, M. Hébert, and G. M. Kelly. Reflective subcategories, localizations and factorization systems. *Journal of the Australian Mathematical Society*, 38:287–329, 1985.

[19] P. Clare, T. Crisp, and N. Higson. Adjoint functors between categories of Hilbert C*-modules. *Journal of the Instititue of Mathematics of Jussieu*, pages 1–33, 2016.

[20] B. Coecke. Terminality implies no-signalling ...and much more than that. *New Generation Computing*, 34(1-2):69–85, 2016.

[21] B. Coecke and A. Kissinger. *Picturing quantum processes: a first course in quantum theory and diagrammatic reasoning*. Cambridge University Press, 2017.

[22] B. Coecke and R. Lal. Causal categories: Relativistically interacting processes. *Foundations of Physics*, 43(4):458–501, 2013.

[23] B. Day. Note on monoidal localisation. *Bulletin of the Australian Mathematical Society*, 8:1–16, 1973.

[24] B. Day. On closed categories of functors II. *Proceedings of the Sydney Category Theory Seminar*, 420:2054, 1974.

[25] F. Elwert. Judea pearl: Causality: Models, reasoning and inference. 2nd edition. *Acta sociologica*, 57:369–371, 2014.

[26] P. Enrique Moliner, C. Heunen, and S. Tull. Space in monoidal categories. In *Quantum Physics and Logic*, volume 266 of *Electronic Proceedings in Theoretical Computer Science*, pages 399–410, 2017.

[27] P. Enrique Moliner, C. Heunen, and S. Tull. Tensor topology. *arXiv:1810.01383*, 2018.

[28] T. Fiore. *Pseudo limits, biadjoints, and pseudo algebras: categorical foundations of conformal field theory*. American Mathematical Society, 2006.

[29] B. Fong and D. Spivak. *Seven Sketches in Compositionality: An Invitation to Applied Category Theory*. Cambridge University Press, 2019.

[30] M. Frisch. *Causal Reasoning in Physics*. Cambridge University Press, 2014.

[31] T. Fritz. Categories of fractions revisited. *arXiv:0803.2587*, 2008.

[32] S. Fujii, S. Katsumata, and P.-A. Melliès. Towards a formal theory of graded monads. In *Foundations of Software Science and Computation Structures*, pages 513–530. Springer, 2015.

[33] P. Gabriel and M. Zisman. *Calculus of fractions and homotopy theory*. Springer, 1967.

[34] J. Goguen. A categorical manifesto. *Mathematical Structures in Computer Science*, volume 1(1):49–67, 1991.

[35] M. Grandis. Cohesive categories and manifolds. *Annali di Matematica Pura ed Applicata*, 157(1):199–244, 1990.

[36] P. Grillet. Regular categories. In *Exact categories and categories of sheaves*, pages 121–222. Springer, 1971.

[37] H. Halvorson, editor. *Deep Beauty : Understanding the Quantum World through Mathematical Innovation*. Cambridge University Press, 2011.

[38] M. Hasegawa. Bialgebras in rel. *Electronic Notes in Theoretical Computer Science*, 265:337–350, 2010.

[39] P. Hayden and A. May. Summoning information in spacetime, or where and when can a qubit be? *Journal of Physics A*, 49(17):175304, 2016.

[40] C. Heunen. An embedding theorem for Hilbert categories. *Theory and Applications of Categories*, 22(13):321–344, 2009.

[41] C. Heunen and M. Karvonen. Monads on dagger categories. *Theory and Applications of Categories*, 31(35):1016–1043, 2016.

[42] C. Heunen and M. L. Reyes. Frobenius structures over Hilbert C*-modules. *Communications in Mathematical Physics*, 361(2):787–824, 2018.

[43] C. Heunen and J. Vicary. *Categories for quantum theory an introduction*. Oxford University Press, 2019.

[44] P. Hines. *Categories and Types in Logic, Language, and Physics*, volume 8222 of *Lecture Notes in Computer Science*, chapter Classical structures based on unitaries, pages 188–210. Springer, 2014.

[45] G. B. Im and G. M. Kelly. A universal property for the convolution monoidal structure. *Journal of Pure and Applied Algebra*, 43:75–88, 1986.

[46] B. Jacobs. *Categorical logic and type theory*. Studies in logic and the foundations of mathematics ; v. 141. Elsevier Science, Amsterdam ; New York, first edition.. edition, 1999.

[47] B. Jacobs and J. Mandemaker. Coreflections in algebraic quantum logic. *Foundations of Physics*, 42(2):932–958, 2012.

[48] P. Johnstone. *Stone spaces*. Cambridge University Press, 1982.

[49] P. T. Johnstone. *Sketches of an elephant: A topos theory compendium*. Oxford University Press, 2002.

[50] A. Joyal. Les theoremes de Chevalley-Tarski et remarques sur l'algebre constructive. *Cahiers de Topologie et Géometrie Differentiele*, XVI-3:256–258, 1975.

[51] A. Joyal, R. Street, and D. Verity. Traced monoidal categories. *Mathematical Proceedings Of The Cambridge Philosophical Society*, 119:447–468, 1996.

[52] I. Kaplansky. Modules over operator algebras. *American Journal of Mathematics*, 75:839–853, 1953.

[53] D. Kartsaklis, S. Pulman M. Sadrzadeh, and B. Coecke. *Reasoning about meaning in natural language with compact closed categories and Frobenius algebras*, page 199–222. Lecture Notes in Logic. Cambridge University Press, 2016.

[54] M. Kashiwara and P. Schapira. *Categories and sheaves*. Springer, 2005.

[55] G. G. Kasparov. Hilbert C*-modules: Theorems of Stinespring and Voiculescu. *Journal of Operator Theory*, 4:133–150, 1980.

[56] A. Kissinger, M. Hoban, and B. Coecke. Equivalence of relativistic causal structure and process terminality. *arXiv:1708.04118*, 2017.

[57] A. Kissinger and S. Uijlen. A categorical semantics for causal structure. In *Logic in Computer Science*, 2017.

[58] J. Kock and W. Pitsch. Hochster duality in derived categories and point-free reconstruction of schemes. *Transactions of the American Mathematical Society*, 369:223–261, 2017.

[59] R. Krömer. *Tool and object: a history and philosophy of category theory*. Birkhäuser, 2007.

[60] A. Kshemkalyani and M. Singhal. *Distributed computing: Principles, algorithms, and systems*. Cambridge University Press, 2008.

[61] R. Lal. *Causal structure in categorical quantum mechanics*. PhD thesis, University of Oxford, 2012.

[62] E. C. Lance. *Hilbert C*-modules: a toolkit for operator algebraists*. Cambridge University Press, 1995.

[63] T. Leinster. *Basic Category Theory*. Cambridge University Press, 2014.

[64] T. Leinster. Rethinking set theory. *The American Mathematical Monthly*, 121(5):403–415, 2014.

[65] D. Lipkind. Russell on the notion of cause. *Canadian Journal of Philosophy*, 9(4):701–720, 1979.

[66] S. Mac Lane. *Categories for the Working Mathematician*. Springer, 2nd edition, 1971.

[67] K. Martin and P. Panangaden. A domain of spacetime intervals in general relativity. *Communications in Mathematical Physics*, 267(3):563–586, 2006.

[68] M. Menni. *Exact completions and toposes*. PhD thesis, University of Edinburgh, 2000.

[69] B. Milewski. *Category Theory for Programmers*. Blurb, 2019.

[70] C. J. Mulvey. Banach sheaves. *Journal of Pure and Applied Algebra*, 17(1):69–83, 1980.

[71] O. Oreshkov, F. Costa, and Č. Brukner. Quantum correlations with no causal order. *Nature Communications*, 3(1):1092, 2012.

[72] R. Penrose. *Techniques in Differential Topology in Relativity*. SIAM, 1972.

[73] J. Picado and A. Pultr. *Frames and locales: topology without points*. Birkhäuser, 2012.

[74] R. Piedeleu, D. Kartsaklis, B. Coecke, and M. Sadrzadeh. Open system categorical quantum semantics in natural language processing. In *Leibniz International Proceedings in Informatics*, volume 35, pages 270–289, 2015.

[75] B. Pierce. *Basic category theory for computer scientists*. Foundations of Computing. MIT Press, 1991.

[76] D. Quillen. Module theory over nonunital rings. `http://www.claymath.org/library/Quillen/Working_papers/quillen%201996/1996-2.pdf`, 1996.

[77] M. A. Rieffel. Induced representations of C*-algebras. *Advances in Mathematics*, 13(2):176–257, 1974.

[78] K. Rosenthal. *Quantales and their applicatoins*. Longman, 1990.

[79] W. Rudin. *Real and complex analysis*. Tata McGraw-Hill Education, 2006.

[80] B. Russell. On the notion of cause. *Proceedings of the Aristotelian Society*, 13:1–26, 1912.

[81] P. Selinger. A survey of graphical languages for monoidal categories. In B. Coecke, editor, *New Structures for Physics*, number 813 in Lecture Notes in Physics, pages 289–356. Springer, 2009.

[82] D. Spivak. *Category Theory for the Sciences*. The MIT Press, Cambridge, 2014.

[83] R. Street. The formal theory of monads. *Journal of Pure and Applied Algebra*, 2(2):149–168, 1972.

[84] K. Szlachányi. Skew-monoidal categories and bialgebroids. *Advances in Mathematics*, 231(3-4):1694–1730, 2012.

[85] S. Thompson. *Type theory and functional programming*. International computer science series. Addison-Wesley, Wokingham, 1991.

[86] Č. Brukner. Quantum causality. *Nature Physics*, 10(4):259, 2014.

[87] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. `https://homotopytypetheory.org/book`, Institute for Advanced Study, 2013.

[88] J. van Benthem and G. Bezhanishvili. *Modal Logics of Space*, pages 217–298. Springer, 2007.

[89] F. Verstraete and H. Verschelde. On quantum channels. *arXiv:quant-ph/0202124*, 2002.

[90] S. Vickers. *Topology via Logic*. Cambridge University Press, 1989.

[91] R. Wald. *General relativity*. The University of Chicago Press., 1984.

[92] G. Winskel. *Events in computation*. PhD thesis, University of Edinburgh, 1980.

[93] G. Winskel. Event structures. In *Lecture Notes in Computer Science*, volume 255, pages 325–392. Springer, 1987.

Index